Reimbursement Strategies for Nursing Home Care:

PROGRAM EXPERIMENTATION
BRANCH LIBRARY

Developmental Cost Studies

Volume II
Technical Papers

Howard Birnbaum
Christine Bishop
Gail Jensen
A. James Lee
Douglas Wilson

June 1979

PROGRAM EXPERIMENTATION
BRANCH LIBRARY

REPORTS
RA
997
R45
1979
v.2

repared By:
bt Associates Inc.
5 Wheeler Street
ambridge, Massachusetts 02138

For:
Health Care Financing Administration
Department of Health, Education and Welfare
Contract No. 600-77-0068

his report is made pursuant to contract 600-77-0068. The amount charged to the Department of Health, Education, and Welfare for the ork resulting in this report (inclusive of the amounts so charged for any prior reports submitted under this contract) is $216,077.

ORDS Research Reference Center

ABT ASSOCIATES INC.
55 WHEELER STREET, CAMBRIDGE, MASSACHUSETTS 02138
TELEPHONE • AREA 617-492-7100
TWX: 710-3201382

RA
997
.R45
1979
v.2

AAI No. 79-63

PROGRAM EXPERIMENTATION
BRANCH LIBRARY

REIMBURSEMENT STRATEGIES FOR NURSING HOME CARE: DEVELOPMENTAL COST STUDIES

Volume II: Technical Papers

June 1979

by: Howard Birnbaum
Christine Bishop
Gail Jensen
A. James Lee
Douglas Wilson

CMS Library
C2-07-13
7500 Security Blvd.
Baltimore, MD 21244

PROGRAM EXPERIMENTATION
BRANCH LIBRARY

This report is made pursuant to Contract No. 600-77-0068 between the Health Care Financing Administration, HEW, and Abt Associates Inc. The amount charged to HCFA for the work resulting in this report (inclusive of the amounts charged for any prior reports submitted under this contract) is $216,077. The names of the persons employed or retained by the contractor with managerial or professional responsibility for such work or for the content of the report are: Howard Birnbaum, Christine Bishop, Gail Jensen, A. James Lee, and Douglas B. Wilson.

The project officer for this contract was Dr. Thomas Kickham, Office of Demonstration and Evaluation, Office of Research, Demonstration, and Statistics.

The views and opinions expressed in this report are the contractor's and no endorsement by the Health Care Financing Administration or Department of Health, Education, and Welfare is intended or inferred.

AN EQUAL OPPORTUNITY EMPLOYER

Table of Contents

		Page
Technical Paper #1:	REVIEW OF THE LITERATURE by Christine Bishop, Patricia Reagan, and Howard Birnbaum	T-1-1
Technical Paper #2:	REIMBURSEMENT SYSTEMS IN OPERATIONS by Christine Bishop	T-2-1
Technical Paper #3:	METHODOLOGICAL AND ECONOMETRIC APPROACH TO COST ANALYSES by Gail Jensen, Howard Birnbaum, Christine Bishop, and A. James Lee	T-3-1
Technical Paper #4:	AN ANALYSIS OF NUSING HOME COSTS BASED ON A NATIONAL SAMPLE by Gail Jensen and Howard Birnbaum	T-4-1
Appendix T-4.A:	DISCUSSION OF DATA	T-4-A-1
Appendix T-4.B:	SUPPLEMENTAL TABLES	T-4-B-1
Appendix T-4.C:	COST COMPONENT ANALYSIS	T-4-C-1
Technical Paper #5:	NURSING HOME COSTS IN NEW YORK by A. James Lee and Howard Binrbaum	T-5-1
Appendix T-5.A:	NEW YORK STATE NURSING HOME REIMBURSEMENT SYSTEM by Michael Sumner	T-5-A-1
Technical Paper #6:	NURSING HOME COSTS IN MASSACHUSETTS by Christine Bishop	T-6-1
Appendix T-6.A:	THE MASSACHUSETTS RATE SETTING SYSTEM	T-6-A-1
Appendix T-6.B:	MEANS AND STANDARD DEVIATIONS OF MASSACHUSETTS DATA	T-6-B-1
Technical Paper #7:	NURSING HOME COSTS IN INDIANA by Gail Jesnen	T-7-1
Appendix T-7.A:	INDIANA STATE NURSING HOME REIMBURSEMENT SYSTEM	T-7-A-1
Appendix T-7.B:	MEANS AND STANDARD DEVIATIONS OF INDIANA DATA	T-7-B-1
Technical Paper #8:	A MULTI-EQUATION MODEL OF NURSING HOME COST BEHAVIOR by A. James Lee, Howard Birnbaum, Christine Bishop, and Gail Jensen	T-8-1

TECHNICAL PAPER #1

Technical Paper #1

REVIEW OF THE LITERATURE*

T-1.0 THE NURSING HOME AS A FIRM: OBJECTIVES AND CONSTRAINTS

The description presented below of nursing homes and their patients and of government intervention in the nursing home industry, will form the basis for our model of nursing home behavior. The behavior of other actors in the system, specifically consumers and regulators, will be seen as constraining the nursing home operator's optimization process. The nursing home as an organization is seen as:

- maximizing an objective function,
- subject to both revenue and technical constraints, including
  - --production technology (a production function),
  - --input costs,
  - --a demand function for the product, including both public and private demand,
  - --government regulations, which affect the technology, cost, and demand constraints.

The form, arguments, and parameters of the objective function and the constraints determine the way nursing homes both as individual firms and as an industry, respond to changes in exogenous and policy variables. These objective and constraint functions will be considered in turn. Evidence from the literature on nursing homes will be used as appropriate. Unfortunately, most analytical students of long term care facilities have not undertaken to construct complete models of the industry's behavior, but rather have focused on certain aspects, like costs or capacity trends. Less analytical studies are also brought to bear on questions about nursing home demand and regulation, for they indicate which variables may be important and at least the direction of gross impacts.

---

*by Christine Bishop, Patricia Reagan, and Howard Birnbaum

T-1.1 Objectives of Nursing Home Operators

The objectives of operators managing nursing homes of differing ownership types may be very different. It is assumed here that for-profit operators attempt to maximize total wealth, though this objective may translate into differing objective functions for different types of non-profit facilities; while operators of non-profit facilities maximize a utility function whose arguments may include output in patient days, quality, number of individuals served, gross revenue and/or value of physical facilities. The for-profit case will be considered first.

T-1.1.1 Objectives of For-Profit Operators

The high proportion of for-profit firms in the nursing home industry has allowed researchers to argue that standard economic models of profit maximization can be transferred to this portion of the health sector. For example, Skinner and Yett (1970) assume cost minimization, p. 3; AMS (1976, pp. 3.55-3.60) develops a model of price discrimination which assumes profit maximization; profit maximization is implicit in the analysis of Friedman (1975).

However, an economic model postulating maximization of the difference between revenues and costs does not fully capture the long run process of wealth maximization for the organizations that operate nursing homes. This is vividly demonstrated in the report of an investigation of proprietary nursing homes in New York State (Moreland Act Commission (1976)); see especially Report 2, Reimbursement of Nursing Home Property Costs: Pruning the Money Tree) and in Shulman and Galanter (1976). Both underscore the value of cash flow to nursing home operators. As in other real estate investments, a small amount of owner's equity can produce a substantial cash flow net of taxes, due to the fact that depreciation is an expense but not a cash outlay, and interest payments are deductible for income tax purposes. Shulman and Galanter do not add, as they might, that when interest and depreciation are fully paid for as an allowable expense by a cost-based reimbursement system, and the stream of patients is much greater than current bed supply

levels, the cash flow looks like much more of a "sure thing" than it would for an apartment, a hotel, or office building which must operate in a competitive market where demand conditions (price and utilization) can change quite rapidly. (It must be noted, however, that nursing home operators do face other kinds of risks, mainly uncertainty about future government action.)

For many operators, then, wealth may be maximized through careful attention to the timing of outlays for expenses and taxes and of gross cash flow due to loans and sale of appreciating assets, as well as gross revenues due to patient care. In some cases the salary paid to the owner/administrator as part of the public rate and his return on equity under Medicaid may figure heavily in his wealth calculation, while in other cases such rewards are dwarfed by the gains to be reaped from cost-free land speculation (since Medicaid pays the taxes and interest) and the use of free funds from depreciation allowances. While the impact of rate regulation on the value of this maximand is significant, it is not all-important, because the revenues for providing services to public patients are only part of the potential return to ownership of a nursing home. Regulation also affects the owner's wealth through rules restricting real estate transactions, which might be used to build up depreciation allowed under rate setting and to realize appreciation of real estate; through rules prescribing minimum amounts for owner equity (i.e., maximum mortgaging) and limiting interest to be reimbursed under rate setting; and through wider public policy with respect to taxation.

In a full study of both capital and operating inputs and costs, it would be necessary to model the proprietary operator's objective in full. Each variable would be subscripted from t = 1 to t = T, where T is the maximum life of the facility, and variables would describe the following flow of funds elements:

- revenues from private patients
- revenues from public patients
- direct operating costs
- loans taken out in year t
- interest paid, based on loans taken out in years 1 through t
- mortgage amortization, also based on previous loans

- taxes paid on net income, where income is net of depreciation and interest
- cash accruing from sale of assets in year t
- return on net after tax cash flow, years 1 through T

A full model could be directed toward finding the optimum amount of leverage and the optimum year in which to sell a facility, under various assumptions about land appreciation and changes in the value of nursing home structures and equipment. It would also model the path of industry growth under current conditions, and in response to changes in regulations.

The focus here must remain far simpler: namely on the choices made about variable inputs, patient mix, occupancy, and private price with two assumptions: 1) capital investment is given, and 2) financing decisions are separable from operating choices. The operator may initially choose firm bedsize simultaneously with desired occupancy and private-public mix, but after this scale decision is made, it is not easily altered. Thus in the short or intermediate run, capital costs are sunk costs and firms may be seen as maximizing the difference between operating revenues and operating costs.

In summary, in the short and intermediate run, when capital investment and financing methods are given, it may be reasonable to assume that operators indeed maximize the difference between revenue from patients and the non-capital costs of providing care. In other words: operators may choose long-run amounts of capital and labor that include more capital than is necessary for production of the amount and level of care they provide, because use of capital in itself has tax and other advantages; but once capital is given, they choose minimum amounts of variable inputs to produce output, subject of course to regulatory constraints. Given this assumption, the problems of growth in the nursing home and efficient size, can be split off from the problem of non-capital input and costs.

T-1.1.2 Objectives of Non-Profit Operators

Non-profit nursing homes are run by religious groups, fraternal organizations, and local governments as well as by non-profit corporations especially constituted to operate nursing homes. These organizations may well have goals of their own beyond providing efficient effective nursing home services. These goals may include in-kind subsidy of members, conversion

of non-members, and providing jobs to constituents. This suggests that different types of non-profit facilities may be operated to maximize differing objective functions. One way to treat this diversity is to assume that the arguments in the objective function are the same across the various types of non-profit firms, but that the values placed on these arguments may differ. The literature on the objectives of non-profit hospitals (see especially Newhouse, 1970 and Lee, 1971) suggests an objective function including both quantity (which may be measured as patient days or admissions, or some combination of these aspects of output) and quality (usually defined as amount and technological sophistication of inputs, for want of patient outcome measures) as arguments in an administrator's utility function. These ideas may be translated to the nursing home situation without the complications introduced in the hospital case by the powerful role of the physician staff (as in Pauly and Redisch, 1973). For example, especially for public nursing homes, size of the organization, in terms of beds, output, and number of employees, may increase the administrator's satisfaction.

T-1.1.3 Comparison of For Profit and Non-Profit Objectives

There can be no straightforward test of hypotheses about differences in the objective functions of for-profit and non-profit nursing homes. One difficulty is that differences in objective functions are confounded with differences in certain constraints, specifically input prices. Non-profit nursing homes may find the cost of capital lower because of the availability of charitable contributions and the absence of local real estate tax and corporate income tax, and may face lower labor costs because of work performed by volunteers and members of religious orders.

However, it will be interesting to compare the behavior of firms by ownership. Some of these differences have been presented and explored in the literature. For example, with respect to the value of scale per se, it is observed that non-profit nursing homes do in fact produce more patient days per facility, and are on average larger in bedsize. (U. S. DHEW,1975)

In addition, non-profit industry spokespeople have argued that quality of care is an argument in their objective function and that they are willing to care for more disabled patients. Holmberg and Anderson (1968) compared non-profit and for-profit facilities in Minnesota and found no significant differences in their indicators of quality (mostly input proportions like staff/patient ratios and patient participation in activities) except physician hours per patient were significantly higher in the non-profit facilities; the only significant difference in an extensive list of patient characteristic comparisons was that proprietary facilities had significantly more mentally confused and disoriented patients. However, such comparisons, while interesting, cannot prove or disprove hypotheses about non-profit and for profit objective functions.

T-1.1.4 Impact of Regulation on Objective Functions

The actual maximand, be it net wealth for proprietary facilities or a function of quantity, quality, and patient mix (for non-profit facilities) will of course not be changed by regulation, except insofar as regulation changes the nature of the organizations providing nursing home services. For example, non-profit facilities could be favored by planning bodies. It is possible that the move to encourage professional training for nursing home operators, through licensing requirements, is basically an attempt to change objectives by changing tastes through education. The assumption appears to be that administrators who have been sensitized to the needs of the elderly are more likely to include quality of care in their objective function. However, reliance on professional ethics of providers to assure quality is not usually considered a realistic approach to the nursing home situation*, in contrast to the case of physicians and hospitals, who perhaps are too often considered by the public to be acting for the public good.

---

*The public apparently is upset that real estate magnates, hotel owners, and other business people have flocked into the nursing home business, while if these people had diversified into supplying car washing services or restaurant meals, no one would mind. The business people are seen as unprepared to provide high quality service, though if it was in their interest to do so, as it would be with car washes and fast food due to market competition, we can be sure that they would learn how, or purchase the necessary technical expertise. There appears to be general suspicion about operators who provide health services for profit.

T-1.2 Production Constraints

The product of nursing homes is multi-dimensional. It is typically measured in patient days, i.e., number of different patients served by the facility in a year times the average number of days each stays within that year. Admissions are possibly a relevant dimension of output, so that, for example, more inputs are required to produce 300 patient days (of a given quality for a given case mix) for ten patients each staying 30 days than for one patient staying 300 days. The number of patient days will be represented by the variable Q and the number of admissions by the variable A. Average length of stay is calculated by $\frac{Q}{A}$, or patient days per admission. Case mix and quality of care are attributes of nursing home care which are very difficult to measure. However, researchers on nursing home costs and production appear to have as an ideal a production relationship Φ with both case mix (X) and quality (q) as arguments:

$$Q = \Phi(K_1 \ldots K_n, L_1 \ldots L_m, x, q, A) \qquad (1)$$

where $K_1 \ldots K_n$ are n types of capital inputs and $L_1, \ldots L_m$ are m type of labor. With q and X constant, a nursing home operator would move along the production function, producing more or fewer patient days of a given quality and for a given case mix, or altering inputs in response to changing prices. Changes in either q or X mean production of days of greater or lesser quality or more difficult or less difficult patient mix. Walsh (1976) distinguishes the impact of quantity and case mix on costs in this way, meaning both are implicitly arguments in his production function. The impact of a shift in q on quantity of output and care type is implicit in work on the hospital sector (see Newhouse, 1970, and Feldstein, 1971, who assume tradeoffs between quality and quantity for a given budget constraint).

T-1.2.1 Case Mix

Patients can be described with respect to medical diagnosis, independence in the activities of daily living, nursing requirements and the like. Assume that each patient j can be evaluated on a set of n dimensions, d, and that a vector $(d_{1j} \ldots d_{nj})$ can be used to describe each patient. Case mix in the facility can be measured by classifying the patients into a number of patient types and calculating the proportion in each type:

$(x_1 \ldots x_m)$. In the analysis below the variable x will stand for this vector describing case mix in a facility. (There is voluminous literature on classification of individual patients and development of debility indices, and some researchers have applied this to facility case mix. See Skinner and Yett (1973) and Winn (1975) for an introduction to this literature as it applies to nursing homes.)

T-1.2.2 Quality

Patient days also have a quality dimension, q. The term "quality" is given many meanings in work by other researchers and it is crucial that it be defined as precisely as possible. Two views of quality, relating quality to inputs and to patient outcome, will be considered. In contrast, the view of quality to be employed in the current model has to do with the adequacy of inputs to deal with case mix.

The first view of quality to be considered is the regulator's view. If regulators were truly specifying a well-defined level of quality, they would constrain the production function ϕ so that

$$q \geq q_r$$

where q_r is the minimum quality level, determined by regulation. However, it is typically assumed that more and better inputs produce high quality as opposed to low quality patient days. (See state regulation of quality, discussed in 3.2.2 and 3.2.3.) Therefore regulation of quality is usually carried out in terms of inputs into production. For different types of capital and labor (e.g., floor space, fire proofing, registered nurses, aides, etc.) the regulations specify that inputs (but not necessarily all inputs) reach minimum levels per patient day or per bed. For example,

$$\frac{K_j}{Q} \geq k_{jr} \quad \text{or} \quad \frac{K_j}{B} \geq k_{jr}$$

$$\frac{L_j}{Q} \geq l_{jr} \quad \text{or} \quad \frac{L_j}{B} \geq l_{jr}$$

Q = patient days

j = types of capital and labor

B = beds

l_r, k_r = constants determined by regulation

For some inputs, an absolute level is required that is not dependent on the size of the facility measured in beds or the actual scale of production Q; examples are requirements that administrators be licensed, that dietician consultants be available, and that the facility have a medical director. This type of requirement can be specified as

$$K_j \geq K_{jr}$$

$$L_j \geq L_{jr}$$

and presumably increases both the minimum outlay necessary for entry and the opportunity for economies of scale.

All of these requirements depend on the level of care for which the facility is licensed, and its certification if any. These regulations may put some kind of floor beneath quality, if the minimum inputs in fact are effectively directed toward patient health. However, this view does not provide a good indicator of quality level above or below the specified minimum; a quality measure based on inputs per patient day implies that twice as much nursing or housekeeping or capital inputs is twice as good, which may not be the case, especially if patient mix can also vary. Moreover, such a measure does not allow tests of hypotheses about efficiency: an increase in inputs per patient day would be interpreted to mean that more quality is being produced, while this could as easily mean that inefficiency has increased.

At the other extreme is a measure of quality based on patient outcome. Under this view, the patient himself is clearly seen as an input into the production process. He enters the facility with certain medical conditions, and in theory a prediction can be made about what his condition will be in three months, six months, or a year without care. The nursing home then goes to work on him, and meets, betters, or falls short of this prediction, thus revealing its quality of care (see Maker, Winn and McCaffree, n.d., pp. 28-31). This view of quality focuses attention where it belongs, on what happens to the patient. However, it requires extensive information on patients at several points in time, and thus is not administratively feasible for incorporation into a rate setting system. The outcome oriented view

points up two important aspects of the term "quality" as it is typically used: first, that quality ultimately is measured by patient outcome, and second, that quality of a given level probably requires differing amounts and mixes of inputs when patient conditions differ. Thus more inputs may mean higher quality, case mix held constant.

However, if case mix should happen to worsen as inputs rise, the increased inputs may not be sufficient even to maintain the previous quality level. In order to capture such a situation, the nursing home may be seen as producing an intermediate good, services, and attaching these to another product, a "stripped down" patient day. Services are of three types:

- nursing care
- personal care
- residential care

Various types of nursing services are performed by nurses, personal care tends to be carried out by aides, and the intensity of residential care (amenities, hotel services, recreation) is a function of capital and housekeeping staff. The concept of patient services is useful in that it allows quality to be a function of the services per patient day with respect to case mix, and allows the adequacy of <u>services</u> rather than quality itself in the production relationship. Thus "quality" depends on inputs, case mix, and the efficiency with which inputs are directed toward production of services to patients. Patients prefer higher quality patient days, insofar as they are able to judge it, so that quality, q, a function of services S and case mix X, appears in the demand function. But the firm produces, not quality, but the intermediate good S.

(2)	$Q = Q(K_1, L_1, A)$	patient day production
(3)	$S = S(K_2, L_2, Q)$	services production, amount per patient day
(4)	$q = q(S, X)$	quality
(5)	$K = K_1 + K_2$	total capital
(6)	$L = L_1 + L_2$	total labor

To repeat, the patient day production relationship does not include quality itself. What is produced is a joint product, patient days with a bundle of services attached to them--and these services, in conjunction with case mix, determine quality. However, it is not possible to distinguish between inputs going into minimal patient day production (K_1, L_1) into additional services (K_2, L_2). If we can observe services themselves, the production function for patient days is

$$Q = \emptyset(K, L, A, S).$$

If instead we have some measure of adequacy of services for the case mix, or quality, we must include both this measure (q) and case mix in the function:

$$Q = \emptyset(K, L, A, \underset{\text{unobserved}}{\underset{\uparrow}{q(X,S)}}, X)$$

T-1.2.3 Quality Production

Of course the quality function q = (X,S) is important in itself, since q will enter the demand function. It is worth considering this relationship in more detail at this point, using an example for the quality production function. Assume a very simple function for quality, where patients of type x_i could be assumed to require Z_i services. The adequacy of services would be shown by

$$Q = S - \Sigma, Z_i x_i$$

or in vector notation

$$Q = S - ZX$$

Figure T-1.1 shows this quality function. It equals zero when S = ZX, so the function cuts the horizontal plane (services are neither less nor more than minimum medical requirements) in a straight line.

The yellow "awning" shows the relationship between case mix, services, and quality. More interesting functions can be envisioned, for example, with a cross section for a given X looking like Figure T-1.2. so that as S approaches S*, these patients simply cannot be cared for,

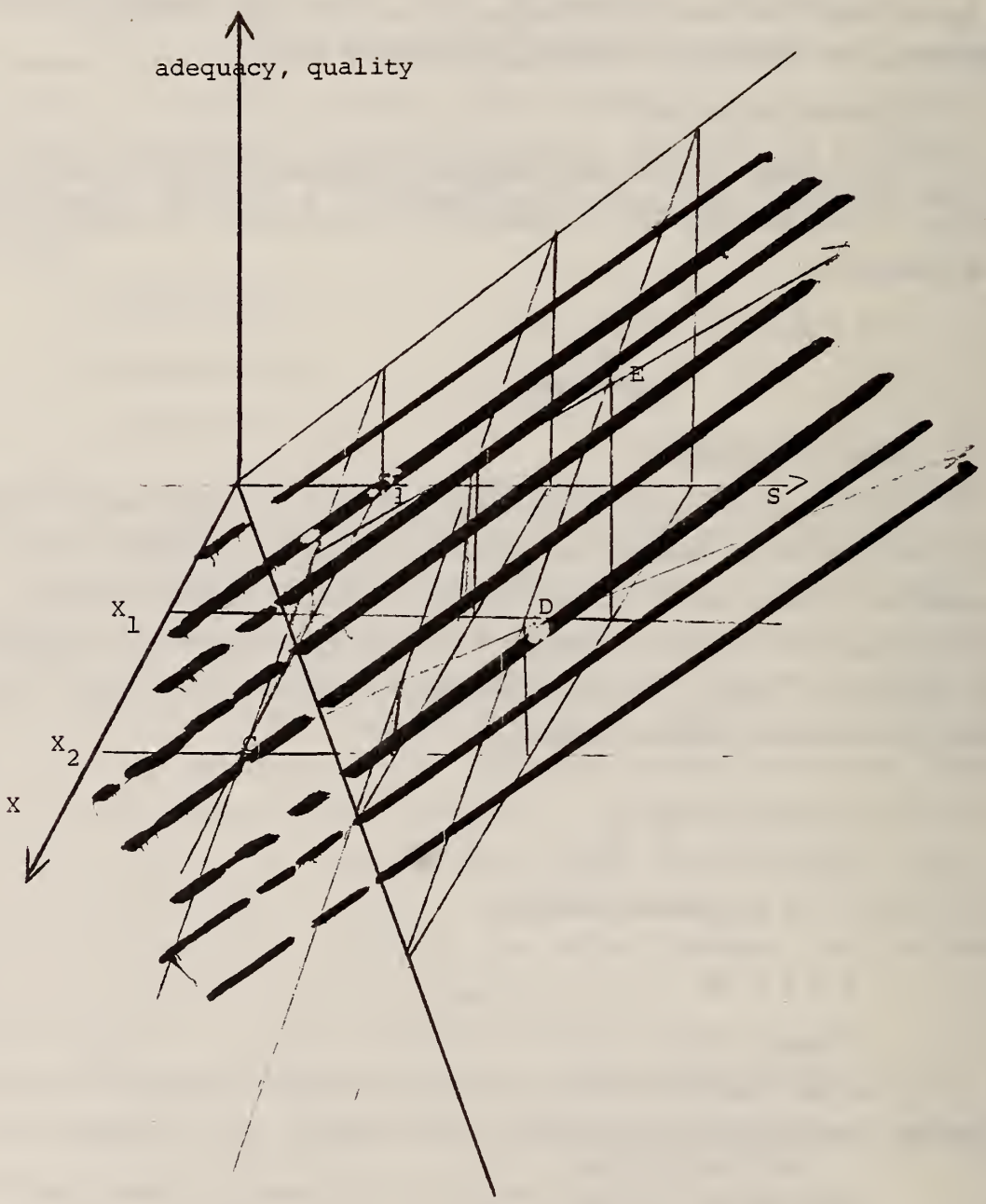


Figure T-1.1

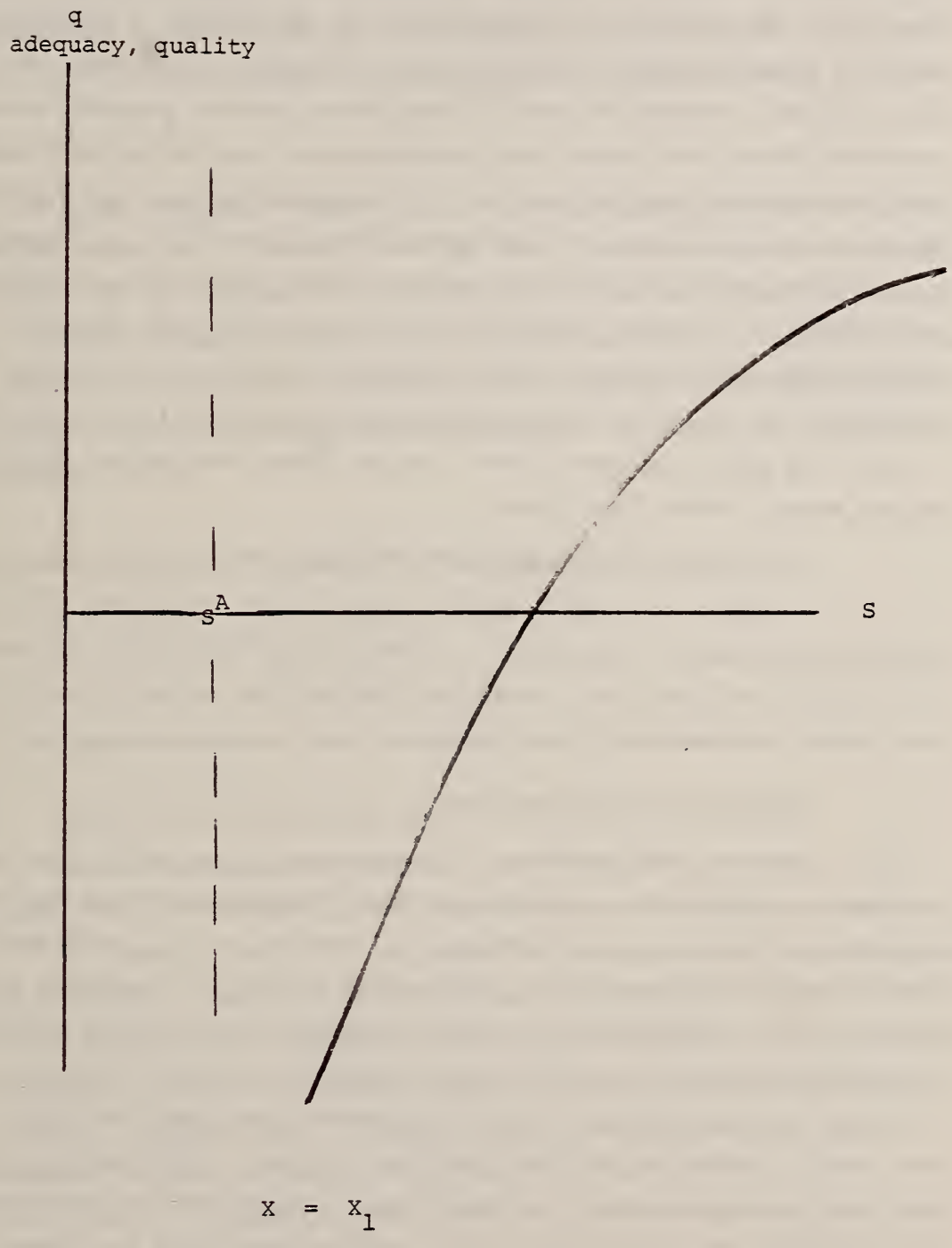


Figure T-1.2

and returns to additional services above the "medically indicated" level where the q index is 0 decrease rapidly.

Potential consumers are more likely to judge quality on their sense of the adequacy of inputs (or services) to deal with the current case mix. The choice of a nursing home is not likely to be made on the basis of expected patient outcome, since potential consumers have even less outcome information than do regulators, and are probably more concerned about the comfort and convenience of day-to-day services than the expected medical outcome. In choosing between two facilities, the home with more patient care and amenity inputs will be chosen over the one with fewer only if the case mix of the first is not so much more difficult that the additional inputs are more than absorbed filling additional needs. Given information about q, as we have defined it ($q = q(S,X)$), private patients will be willing to pay for q, and when public patients have a choice among facilities, they will choose more q rather than less.

In summary, in equations (2) through (6), we can observe only Q, A, K, L, and X, returning us to a formulation that looks like (1). Proxies for S may be available: personnel time spent with patients, availability of aide time, therapies, recreation, square feet of common-room space per bed, etc., but these may not be very satisfying.

T-1.2.4 Regulation and Production

Case mix and services, the determinants of quality are imperfectly subject to regulation. Regulations that attempt to control quality by specifying inputs are discussed above. Patient mix is regulated through initial patient placement and utilization review, for Medicare and Medicaid patient only. Monitoring of patient placement acts on each individual placement decision, so that the characteristics of each patient in the facility must be such that they fall between the lower and upper bounds for the level of care the facility provides for him. Note that mix in the facility is not monitored, so that even if every patient individually meets the criteria, the mix for the whole nursing home could be skewed toward the low or the high end of nursing needs.

Further regulations apply to capital, K, through determination of need (DON). These regulations make it difficult to alter capital inputs in response to changing conditions or production methods. Although under DON laws approval is necessary for diminishing capacity, in practice these regulations have their impact on capital increases. This means that a facility's capital endowment at time t must meet the condition that future capital can be no greater than capital in year T.

$$K_t \leq K_T$$

where T is some (probably recent) time. If they are effective in slowing facilities' responses to conditions that would otherwise call for increases in bed capacity and equipment, certificate of need laws could conceivably prevent us from assuming that a long run cost function can be observed in cross section. DON may also be effective in slowing or precluding entry of new facilities, increasing the monopoly power of existing procedures (See Salkever, 1976).

Production, then, follows the relationship

$$Q = \Phi \left[K, L, A, S \right], \text{ or}$$

$$Q = \Phi \left[K, L, A, q(S, X), X \right]$$

subject to regulations of K and L, and certain limits on X.

T-1.2.5 Production in the Literature

Muller, Johnson, and McCaffree (1975) make an exciting start toward specifying and fitting production functions for nursing homes. The outputs of the nursing home are delineated by seven patient day categories: ECF care provided to Medicare beneficiaries, three levels of care provided to Medicaid patients, and the same three levels supplied to private patients. It is important to note that these output definitions allow different types of care to be provided to public and private patients at the same level within one facility, and do not take quality of care or actual case mix (by diagnosis, debility level, or even age and sex) into

account. In contrast, it is assumed in the current modeling effort that patients at a certain licensed level may receive different amounts of care (outputs) in different facilities, due to quality variations, but that once admitted to a facility, private and public patients receive the same care; and that the services provided to patients (S in our notation, standard care units in Muller et al.) represent differing qualities of care according to patient characteristics or needs.

The preferred form of the production relationship shows firms producing an intermediate output, standard care units, using four inputs: nursing hours of registered nurses, licensed practical nurses, and nurse aides, and dietary labor hours. The intermediate good is then distributed to the seven classes of patients. The estimation technique is canonical correlation, allowing the seven dependent output variables to be related to the form input variables.

Muller et al. use an unusual division between fixed and variable inputs: they define fixed inputs as those which are determined by the number of licensed beds, but then include in fixed inputs food, laundry, and housekeeping labor, which presumably can vary with output, i.e., with the number of patients actually served (see p. 28ff.). The implied distinction, never explicitly stated, is between inputs used to produce a day of room and board, and inputs like nursing and food service labor, which can be varied by the firm to produce a more or less service intensive patient day. This is not the conventional distinction between fixed and variable inputs but is interesting nonetheless, and actually coincides to some extent with our hypothetical distinction between Q (patient days) and S (services per patient day).

The production function for standard care units is assumed to be linear; surprisingly, LPN hours are found to contribute less to production of care than nurse aide hours, and dietary hours are insignificant. It must be noted that this specification does not allow for differences in the amount of fixed factors employed, i.e., there is no substitution between capital and labor. Economies of scale are not discussed, and cannot be examined using this formulation. The coefficients of the dependent

variables show the number of standard care units provided to each type of patient; private patients are found to receive more care than public patients with the same classification, and patients in higher classifications receive more care than those in lower classifications. Never considered is the possibility that, as argued below in Section T-1.4 (Optimization), operators make simultaneous decisions about the number of private and public patients to serve and the amount of nursing services to provide, since additional services allow them to charge a higher price to private patients. The results are compatible with our expectations that firms which find it worthwhile to provide more nursing services to all patients will have a higher than average proportion of private patients. This could account for the difference in the coefficients.

While an interesting exploration of nursing home production relationships, the findings of Mueller et al. cannot be used to assess efficiency and the like because of these difficulties.

T-1.2.6 Costs

The total and average cost of nursing home care, with a given level of service and a given rate of admissions, is the result of the firm's optimization process, i.e., maximization of its objective function. Total cost is found by taking the firm's optimum amounts of capital and labor from the production function and multiplying these by their input prices. This is shown by the accounting identity

$$TC = \sum_{j}^{n} r_j K_j + \sum_{j}^{m} w_j L_j$$

where

TC = total costs

r_j = cost of capital of type j

K_j = amount of capital

w_j = wage of labor of type j

L_j = amount of labor of type j

n = number of types of capital

m = number of types of labor

Cost functions will be dealt with in the Section T-1.4 (Optimization) since they are a result of the optimization process. They are discussed here in order to note that input prices form another set of constraints on the firm.

It should be recognized that regulation directed at problems other than nursing home costs have worked to some extent to control input prices in nursing homes, most notably under the Economic Stabilization Program. Government action with more specific impact on nursing home input prices are the extension of the National Labor Relations Act to non-profit voluntary health care providers in 1973, allowing unionization of non-profit facilities; extension of wage and hours regulation to non-profit voluntary nursing homes, phased in from 1967 through 1969; and action in some states to provide low cost capital, again to voluntary health facilities, through finance authorities that sell state-backed bonds.

In addition, certain regulations of input prices have been developed in response to action by proprietary facilities that have been seen by legislators, regulators, and the public as abuse, if not actual fraud. For example, a nursing home operator who leases his facility must usually show that the rent he pays is consistent with actual property cost and/or that he has no interest in the ownership of the facility which might prevent a true "arm's-length" transaction from taking place. Such regulation is designed to insure that input prices are fair market prices or at least approximate market prices. Other abuses have concerned kickbacks required of pharmacists supplying drugs and medical supplies or of suppliers of purchased services. (In a sense, these are or were hidden ways of making money for the operator and so should appear in the objective function -- leading to a situation where, for example, drugs are over-purchased while the actual input price faced by the firm is the market price -- so this kind of thing is not really regulation of input prices faced by the operator.)

T-1.3 Demand Constraints

T-1.3.1 Demand of Individuals for Nursing Home Care

T-1.3.1.1 Theory of Consumer Demand

It is worth returning to the demands of individual consumers to understand the demand constraint faced by a firm providing nursing home care. An individual consumer may be seen as maximizing a personal utility function, whose arguments include health and well-being as well as various goods, subject to his budget constraint and subject to a personal production function for health. Admission to a nursing home and number of patient days consumed given admission are arguments in the individual's production function for health and well-being, as are substitutes along the continuum of care, from hospital services to home care and congregate living arrangements. A given amount of nursing home care will have differing impacts on the health and well-being of an individual with a given stock of health, depending on its quality. Individuals may be assumed to have varying tastes for services of three basic types,

- nursing care
- personal care
- residential care

which are seen as "attached" to each patient day. These services would have differing value to people with differing health and social needs.

The individual's budget constraint may be shifted out by public payment programs, but only if he meets eligibility requirements and chooses publicly certified care.

An individual's demand for nursing home services of a given type (i.e., a given service mix) is thus a function of his health and social needs, income and wealth, price of care to him (which depends on his eligibility for public payment programs), and the prices (and availability) of substitutes for nursing home care, like hospital services, home health care, and the services of family members. Maximizing utility given the resources available to him, including entitlement to public care, the patient may feel compelled to choose nursing home care if substitutes are not available, or are not covered by public program; or to choose care of

a lower quality level than he would be willing to pay for if he is able to seek care only in facilities that agree to take public reimbursement as their full price.

T-1.3.1.2 Information and Agency Relationships

The consumer may be hindered in maximizing utility by lack of information about alternative nursing homes and be unable, as a lay person, either to assess the impact on health of a given service level, adjusted for case mix (i.e., to know his health production function) or to judge quality (as defined here) before he enters a facility. Limitations on search activity may be increased by physical and psychological disability. Under such circumstances the patient's choices about consumption may be made by an agent--a family member, social worker, or a physician prescribing nursing home care. However, perhaps because nursing home care is usually a last resort, patients, their families, and even professionals do not appear to be informed about what constitutes good care, what its health benefits are, and where it can be purchased. Without this information, they may assess the cost of good care as greater than its benefit, or may stop searching too soon and settle for merely adequate care. Thus, the nursing home administrator who attempts to provide better, more effective services may not be rewarded by higher demand in the marketplace. Since consumers make their choice about which facility to enter at one point in time, and often do not have many opportunities to change facilities even if they are dissatisfied, the administrator has incentives to make his home attractive to shoppers for care rather than to provide high quality care for actual residents.

Physicians, who play key roles in determining the demand for hospital care, are noticeable by their absence from the nursing home scene. The conventional wisdom has it that the nursing home patient is not interesting enough or lucrative enough to induce the physician to provide on-going care. It appears that a physician may care for an older person for a good part of his life and then essentially abandon him to the care of the nursing home. This is even more likely in cases where a specialist, who sees his role as dealing with a particular type of illness, rather than the whole patient, cares for the patient during a hospital stay.

T-1.3.1.3 Demand Studies

Actual consumer behavior should be studied using observations on individual choices to estimate demand for nursing home care. Lacking such a demand study, we must turn to consumer surveys of patients already in nursing homes (Baumgarten, Choudhury, Rao, and Ring, 1972; Baumgarten, Rao and Ring, 1975). They show the lack of search behavior in the marketplace: for 78 percent of the nursing home patients surveyed, no other home was considered. Information gathered on the home and criteria for choice were general "favorable impression of facilities" (21%), advice of others (13%), religious affiliation (10%), location (7%), medical care (3%), and cost (2%); 36 percent of the patients either were not involved in the choice of home or didn't know what criteria were used.

The fact that a consumer can find only certain combinations of services in the marketplace makes choice all the more limited. Baumgarten et al. found that

> despite the seemingly large number of nursing homes available,...the nature of their offerings is so segmented that for most families very few homes are suitable (1975, p. 2).

They did not explicitly recognize the role of facility case mix as well as facility services; the purchaser of an automobile receives all its desirable attributes (capacity, horsepower, gas mileage) himself, but when a consumer buys a nursing home day with an average of x hours of nursing or personal care, and z residential amenities per patient, he must share these with others whose health may be such that they absorb much more or much less than their "share." (See Winn, Bennett, and McCaffree, 1976.) For this reason we emphasize "quality," or the adequacy of services in relation to case mix needs.

T-1.3.4 Market Demand for Nursing Home Care

In economic theory, individual demands for a good, here admission and/or patient days, sum to a market demand curve. This implies that the total quantity demanded depends on aggregate versions of some variable that determine the quality demanded by individuals. This quantity of admissions and/or patient days is expected to be a function of price of care, quality of care, income of the population, demographic variables like age and marital status which indicate health and social needs,

(ideally a variable for health status itself), and the prices (and availability) of substitutes and complements.

$$Q_D = Q_D(P, q, Y, d_1 \ldots d_n, p_1 \ldots p_n)$$

where

Q_D = quantity of patient days demanded

P = price of nursing home care

q = quality of nursing home care

Y = income

$(d_1 \ldots d_n)$ = demographic descriptors

$(p_1 \ldots p_n)$ = prices of substitutes and complements

Note that quality appears on the right-hand side of this relationship, in recognition of the fact that nursing home care is a multi-attribute good: more patient days will be demanded ceteris paribus if quality rises and less if quality falls.

As discussed above, the demand for nursing home care is segmented: private patients face a price for care while those eligible for Medicaid face a zero price, and cannot make side payments to facilities. Demand for private days (given quality) is somewhat elastic with respect to price (see Chiswick 1976) but demand for public days by public patients is not responsive to price. Figure T-1.3 shows the public and private demand, and their horizontal sum, the total market demand.

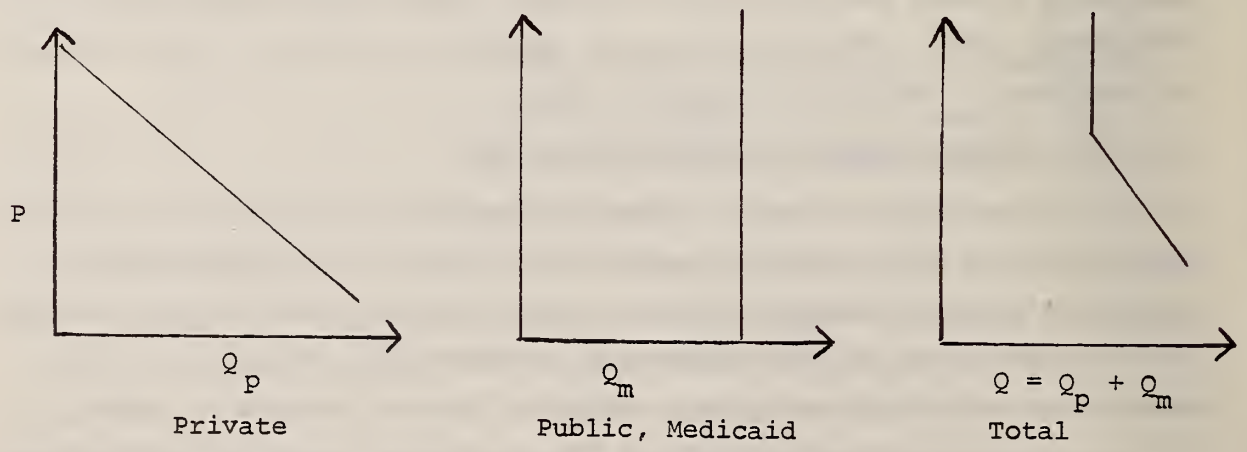


Figure T-1.3

Although the demand of public patients is not restricted by price, other mechanisms (such as waiting lists) restrict public patients' access to care. Nevertheless, evidence indicates that the final outcome is chronic excess demand in the public segment of the market. (See Beattie and Jordan, 1975.)

Chiswick's (1976) study of nursing home aggregate demand can shed some light on market demand. The study uses a cross section of SMSAs for 1967, and the dependent variable is nursing home residents per population sixty-five and over. Demand is seen to be a function of demographic variables (age, sex, mortality rate) which indicate health status; the availability of one substitute, care by family members, proxied by the labor force participation rate of married women; income; a variable to indicate state participation in Medicaid in 1967; and an estimated price variable. Chiswick makes some strong assumptions in order to estimate the price variable; he assumes that the supply curve for nursing home services is perfectly elastic and that its level is determined by the level of various input costs across SMSAs, specifically wages of RNs, cleaning personnel, aides, and availability of physicians, claimed to represent a source of nursing home entrepreneurial talent. The price equation explains only 21 percent of the price variation across SMSAs until a "quality" variable, employees per nursing home resident, is added to the equation; this raises the R^2 to .58. The estimated price variables used in the demand estimation put actual SMSA values into this second equation, but use a mean for the "quality" variable. The price variable thus is meant to represent the price of "standard" care.

Quantity demanded is found to be very responsive to price (elasticity is -2.3), income (.6 to .9), and labor force participation of adult married women (1.4). Demographic variables are significant and have reasonable signs. The Medicaid dummy is insignificant.

A major flaw in this study is its failure to take into account, aside from a dummy variable, the segmentation of the market. Even in 1967, very early in the history of Medicare and Medicaid, private expenditures were only 48.2 percent of expenditures on nursing homes and related care facilities. Chiswick assumes that public agencies that pay for care and individuals eligible for public sponsorship respond to price in exactly the same fashion as private individual demanders: in other words,

where price is high, that agencies do their best to find alternatives to nursing home care for elderly clients. At the very least, explicit recognition should be made of the lower net price paid by recipients both of Medicare and of preMedicaid governmental subsidy for nursing home care, and the zero price faced by Medicaid recipients.

Another study of nursing home demand deals with only a small segment of the market, Medicare. Feldstein (1971) estimates a demand relationship for a particular segment of the market, Medicare, using state data for Fiscal Years 1967 and 1968. Extended care facility (ECF) admissions per Medicare hospital admission are shown to vary positively with ECF beds per enrollee, with private practice physicians per capita, and with the percentage of the population that is white; and negatively with short term general hospital beds per capita. This estimated relationship implies that realized demand for admission was closely and significantly related to availability of supply at least in the early years of the Medicare program.

T-1.3.3 Firm Demand for Nursing Home Care

Given a market demand relationship for nursing home services, what does the demand facing any given nursing home look like? In a competitive market with many sellers, a standard product, and perfect information, the market supply and demand curves determine the price for care, and individual producers must be "price-takers." This means that the demand curve facing the individual firm is perfectly elastic at the market price, and the firm can sell as much as it chooses at the going price. In the nursing home industry there are indeed many sellers, but the product of nursing homes is highly differentiated, both by the amount and mix of quality (services given patient mix) and by geographical location; in addition, information is poor and search is limited. Thus the demand curves showing admissions or patient days demanded by private patients are probably downward sloping to the right. The quantity demanded by public patients is not responsive to the price the firm charges, since the pay nothing for care. Its location at Q_o depends on other attributes of care, for example quality and location. Although Q_o public patients want care, since the Medicaid program pays for care at a rate R_i (specific to firm i) the demand curve experienced by the firm is shown by the dotted

line in Figure T-1.4B: the firm can produce as many public days as it wishes at a price equal to R_i, up to Q_m^o.

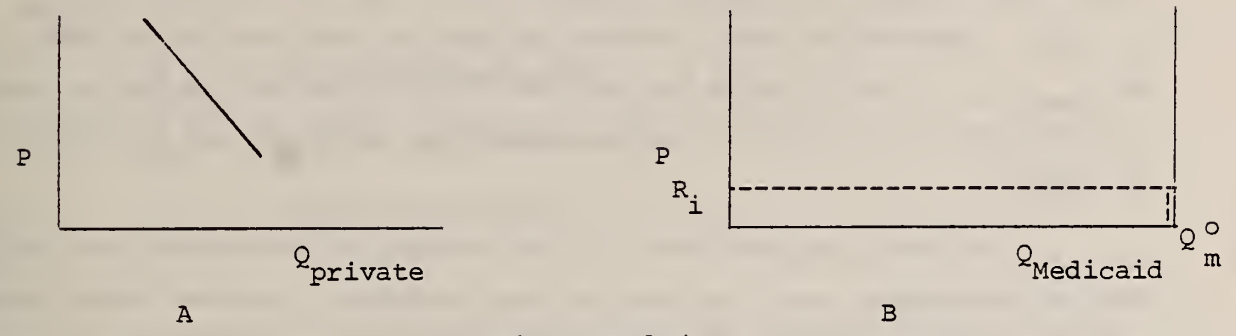


Figure T-1.4

The firm can charge different prices to private and public patients. Assume first that there is no price discrimination among private patients. Then the relevant marginal revenue curve is shown by the dashed line in Figure T-1.5; Q_p private patients are charged P_p, and the state pays R; in the for-profit firm, quantity will be determined by the intersection of the marginal cost curve (MC) with the marginal revenue curve, MR. Once total days to be supplied are determined, the division between public and private patients is also known. For example, if total patient days was Q_T, Q_p private days and $Q_T - Q_p$ Medicaid days would be supplied. The

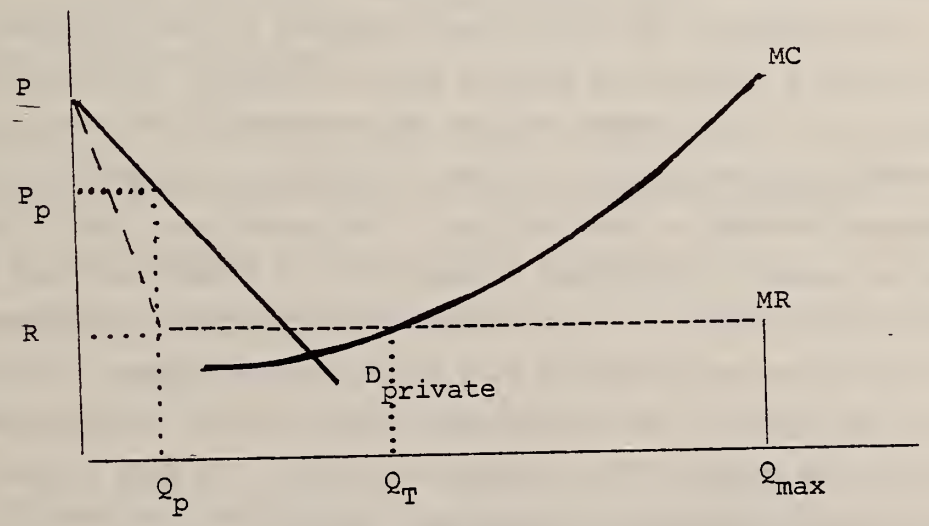


Figure T-1.5

total demand for public days is Q_m^o, from Figure T-1.4B. The maximum quantity demanded of the firm depends on the private quantity, Q_p, the firm chooses; Q_{max} increases as price falls, since it is equal to $Q_p + Q_o$. Unsatisfied public demand for care in this particular home is $Q_{max} - Q_T$. The MC curve can cut the MR curve so that no public days are produced, or so that all public demand is satisfied and some beds are empty.

Medicare patients form a third segment of the market, for skilled nursing facilities only, but are not very numerous. Another complication of market segmentation is that patients may move from private pay status to Medicaid when their funds run out, or from Medicare to Medicaid when covered days are exhausted. It is possible that private prices are set with this in mind.

The private demand curve may be even more complex than has been described here, inasmuch as another attribute of nursing home care is the characteristics of one's fellow patients. Private patients may have a preference for nursing homes that do not serve Medicaid patients, or serve only a small number. This apparent externality in consumption may explain why some facilities with a high proportion of private pay patients have low occupancy rates, even though they could fill up with public patients whose marginal revenue (public price) is presumably greater than marginal cost in a fully staffed facility with empty beds.

The quantity of private days demanded at the individual firm level is thus a function of private price, "quality" (a function of services and case mix), local income, prices and availability of substitutes and complements (including care in other facilities), and percent or number of Medicaid patients in the facility. The quantity of public days demanded depends on quality, prices and availability of substitutes and complements, and rules which determine eligible population; in most circumstances there will be excess demand in this public market segment. The public price is the rate set by the Medicaid agency, which, as described above, depends on the costs to the individual facility, in many states. This makes the rate endogenous to the firm: under most rate setting systems, the nursing home can affect its own price for public care, within limits. This is discussed in Section T-1.4 below.

$$Q_p = Q_p(P, q(S,X), Y, d_1 \ldots d_n, P_1 \ldots P_n, Q_m)$$

$$Q_m^o = Q_m(q,(S,X), d_1 \ldots d_n, P_1 \ldots P_n, E)$$

where

Q_p = private days demanded

Q_m = Medicaid days demanded

E = eligible population

In summary, nursing homes are seen to maximize either the net present value of after tax profits or a utility function including quantity and quality of services; subject to a production function showing the transformation of capital and labor inputs into patient days, adjusted for quality and admissions; subject to input prices; and subject to demand for patient days.

We have discussed the important role played by regulatory activity in each of these constraints, and have attempted to be as specific as possible about the way regulation works in the context of our model.

From the firm's optimization process will emerge in the short run:

- the average cost of care
- the mix of private and public patients
- the private price
- the rate, insofar as it is cost based

In the long run, this optimization by many firms will determine

- the capacity of the industry to supply care.

We turn now to the maximization process.

T-1.4 Optimization by the Nursing Home as Firm

Economists find it useful to assume that the observed behavior of firms is the result of a process in which firm decision-makers maximize their objectives subject to constraints. The discussion above has presented descriptions of these objectives and constraints; it now remains to present the outcomes of optimization process, and to determine whether the postulates about the objectives and constraints facing nursing home operators are in fact helpful in explaining outcomes of the decision-making process, and, if so, whether the model can be used to develop hypotheses with respect to changes in constraints, particularly rate setting methods.

The outcomes observed in the nursing home industry can be viewed in the context of optimization by firms. Researchers investigating the nursing home industry have made a beginning toward describing nursing home behavior, albeit without underlying models to provide explanations of this behavior (see Gottesman, 1974; Greenwald and Linn, 1971; Robinson, Bronfin, and Balswick, 1971; Friedman, 1975; Zimmer, 1975; Curry and Ratliff, 1973). Statistics for nursing homes in particular states (see for example Bishop, 1976; Touche-Ross, 1974) and for the nation (USDHEW 1976) also reveal the end result of the optimization process.

The following situations are observed in the nursing home industry:

1. Nursing homes vary widely in their costs of care.

2. Some nursing homes do not maintain occupancy rates at their maximum, even though there are waiting lists of hospitalized public patients seeking nursing home placement.

3. Facilities vary in their mix of private and public patients, and some clearly cater to either private or public-pay patients.

4. Some facilities provide more nursing hours and other services than the minimum required by regulations.

5. Cost function estimation for nursing homes has shown some impact of case mix on cost, but this relationship is not strong.

6. Nonprofit facilities are on average more costly, larger, and use more labor per patient day.

These issues will be examined more fully in the course of the statistical work in this study, which will make use of both national and state information on individual facilities. The model presented below provides some insight into how these outcomes may arise. Optimization by for-profit nursing home operators will be presented first, followed by a discussion of non-profit optimization

T-1.4.1 For Profit Nursing Home Decision-making

In Section T-1.4 above, it is assumed that the objective of the for-profit operator is to maximize the net present value of after tax return. In the short run, given capital (namely bedsize and equipment) and past decisions about financing, and if it is worthwhile for the facility to operate at all, the operator will maximize the difference between operating revenue from patient care and operating expenditures. Thus it can be assumed that in the short run operating homes maximize revenue minus cost. This assumption does not allow conclusions about optimal capital or optimal financing arrangements, which depend on many other variables in addition to what we will call operating profits. This distinction is useful in terms of developing a prospective reimbursement system for operating, as opposed to fixed, costs, since it allows us to refrain from investigation of capital costs and returns. These costs and returns must be considered if we are to assess viability of the industry and expansion under various PR schemes. Therefore, given capital, the operator will maximize

$$\pi = PQ_p + RQ_m - \sum_j w_j L_j$$

where P = private price

Q_p = private days supplied

R = Medicaid rate

Q_m = Medicaid days supplied

L_j = variable inputs, including labor and suplies

w_j = prices of variable inputs

The constraints faced by the operator are:

- a production function
- a demand function for private days
- a rate setting function
- an accounting identity stating that total days is the sum of public and private days, $Q = Q_p + Q_m$
- a constraint on total patient days, $Q \leq B^* 365$

  where B = beds, a given

The production relationship postulated above is

$$Q = \emptyset\ (K, L, S, A)$$

$$Q = Q\ (K_1, L_1, A)$$

$$K = K_1 + K_2$$

$$L = L_1 + L_2$$

$$S = S(K_2, L_2)$$

The demand relationship for private days is

$$Q_p = Q_p \left[q\ (S, X),\ P,\ Q_m \right]$$

where q = quality, a function of services and case mix

x = case mix

The rate setting function depends on how endogenous the rate process is. For example, under full reimbursement of all costs regardless of occupancy rate,

$$R = \frac{FC + wl}{Q_p + Q_m}$$

where FC = fixed costs, constant for the time period

If occupancy rate is required to reach minimum level $\bar{u}$,

$$R = \frac{FC + wL}{\bar{u}^*B^*365} \quad \text{when} \quad \frac{Q_m + Q_p}{B^*365} < \bar{u}$$

$$R = \frac{FC + wL}{Q_m + Q_p} \quad \text{when} \quad \frac{Q_m + Q_p}{B^*365} \geq \bar{u}$$

where B = capacity measured in beds

If certain costs are not allowed to exceed specific ceilings,

$$R = \frac{FC + \sum_i \min (w_i L_i, C_i)}{Q_m + Q_p}$$

where C_i = upper limits on costs for variable inputs

Elements of fixed costs may also be disallowed, changing the constant FC, to FC_r, fixed costs allowed under rate regulation.

If the rate is set prospectively, based on past costs, the problem becomes dynamic. As an example,

$$R_t = \left(\frac{FC_{t-1} + w_{t-1} L_{t-1}}{B_t * 365 * \hat{u}_t} \right) \quad \frac{\hat{CPI}_t}{CPI_{t-1}}$$

where $\hat{u}$ = expected occupancy rate for year t

$\hat{CPI}$ = expected level of the consumer price index for year t

CPI_{t-1} = actual level of the CPI in the previous year

The impact of various rate setting methods on firm behavior will be interesting to explore using this model. However, for the present, let the rate be set in advance for the firm (i.e., R is neither endogenous nor endogenous with a lag).

R = constant

We also make the unrealistic simplifying assumption that length of stay $(\frac{Q_m + Q_p}{A})$ and case mix (X) are given, so that when Q and S are known, A and q are determined. The operator can vary inputs (L), private price (P), Medicaid patient days (Q_m), and services (S); in this way he determines private days demanded (and thus total patient days) and quality (q = q(S,X)).

Total patient days must be less than or equal to the number of beds x 365. We ignore this for the present, so that the following argument shows <u>desired</u> level for the variables. If $Q_p + Q_m < B*365$, the desired levels will be achieved at a level of occupancy less than 100 percent; if not, the

facility will be full, and the optimization process can be resolved using the identity $Q_p + Q_m = B*365$ to find the mix of private and public days. (A more elegant mathematical formulation would incorporate the inequality explicitly.)

The optimization process can be shown using the Lagrangian:

$$L^* = P*Q_p\left[P,\ Q_m,\ q(S,X)\right] + R*Q_m - wL - rK + \lambda\left[\phi(K,L,S,A) - Q_p - Q_m\right] \quad (1)$$

$$\frac{\partial L^*}{\partial P} = Q_p + P\frac{\partial Q_p}{\partial P} - \lambda\frac{\partial Q_p}{\partial Q_m} - \lambda = 0 \quad (2)$$

$$\frac{\partial L^*}{\partial L} = -w + \lambda\frac{\partial \phi}{\partial L} = 0 \quad (3)$$

$$\frac{\partial L^*}{\partial Q_m} = P\frac{\partial Q_p}{\partial Q_m} + R - \lambda\frac{\partial Q_p}{\partial Q_m} - \lambda = 0 \quad (4)$$

$$\frac{\partial L^*}{\partial S} = P\frac{\partial Q_p}{\partial q}\cdot\frac{\partial q}{\partial S} + \lambda\frac{\partial \phi}{\partial S} - \lambda\frac{\partial Q_p}{\partial q}\frac{\partial q}{\partial S} = 0 \quad (5)$$

From (2) $Q_p + P\dfrac{\partial Q_p}{\partial P} = \lambda\dfrac{\partial Q_p}{\partial P}$

dividing by $\dfrac{\partial Q_p}{\partial P}$

$$Q_p\frac{\partial P}{\partial Q_p} + P = \lambda \quad (6)$$

The left hand side is the marginal revenue due to an additional private patient.

From (3)
$$w = \lambda\frac{\partial \phi}{\partial L} \quad (7)$$

In other words, the wage equals the marginal product of labor.

Form (4)
$$R = (\lambda - p)\frac{\partial Q_p}{\partial Q_m} + \lambda \quad (8)$$

Substituting from (6)

$$R = \underset{(-)}{Q_p\frac{\partial P}{\partial Q_p}}\cdot\underset{(-)}{\frac{\partial Q_p}{\partial Q_m}} + \text{marginal revenue of an additional private patient}$$

Private and public patient days will be produced so that the rate is equal to the marginal revenue of an additional private patient plus the impact (negative) on private revenue due to increasing public patient day

production. The private demand curve moves leftward with every additional public patient, making this impossible to draw in two dimensions.

Substituting (6) into (5)

$$P\frac{\partial Q_p \partial q}{\partial q \partial S} - \frac{\partial Q_p \partial q}{\partial q \partial s}\left[Q_p \frac{\partial p}{\partial Q_p} + p\right] = -\frac{\partial \phi}{\partial S}\lambda$$

$$\underset{(-)}{Q_p}\ \underset{(-)}{\frac{\partial P}{\partial Q_p}}\ \underset{(+)}{\frac{\partial Q_p}{\partial q}}\ \underset{(+)}{\frac{\partial q}{\partial S}} = \underset{(-)}{\frac{\partial \phi}{\partial S}} * \lambda \qquad (9)$$

The left-hand side shows the revenue gain for a marginal increase in service production due to the fact that when services increase, prices can rise without a reduction in Q_p; the right hand side shows the revenue loss due to the fact that an increase in services production entails a shift of resources from production of patient days (with value λ) to services at the rate of $\frac{\partial \phi}{\partial S}$.

If the rate R increases, equation (8) implies that private patient days will fall, since the rate will equal the marginal revenue of a private patient at a lower value for Q_p.

If case mix worsens, assuming that the quality function declines as case mix worsens ($\frac{\partial Q}{\partial X} < 0$,) the facility will become less desirable to private patients (demand curve shifts leftward); the optimal level of services may well increase, shifting the private demand curve rightward a bit--but in all probability the increase in S will not fully compensate for the increase in X, so that quality will fail.

In the process of maximizing operating profits, firms are likely to reach different decisions about occupancy rates, mix of public and private patients, service production, and average cost. This happens for two reasons. First, with bedsize B fixed in the short run, and unable because of certificate of need regulation and capital costs to adjust rapidly, facilities will have differing values for B in relation to their desired Q. Secondly, it can be argued that firms start out with <u>different private demand curves</u> due to facility characteristics, like age of facility, physical plant, and reputation, that cannot be easily altered in the short run. Thus one firm may find it worthwhile to add services, increasing quality to attract private patients; while another, with low private demand, would not be able to increase private patients above zero without such a massive increase in S that this cannot be worthwhile. This affects the firm's choice of patient mix and average cost.

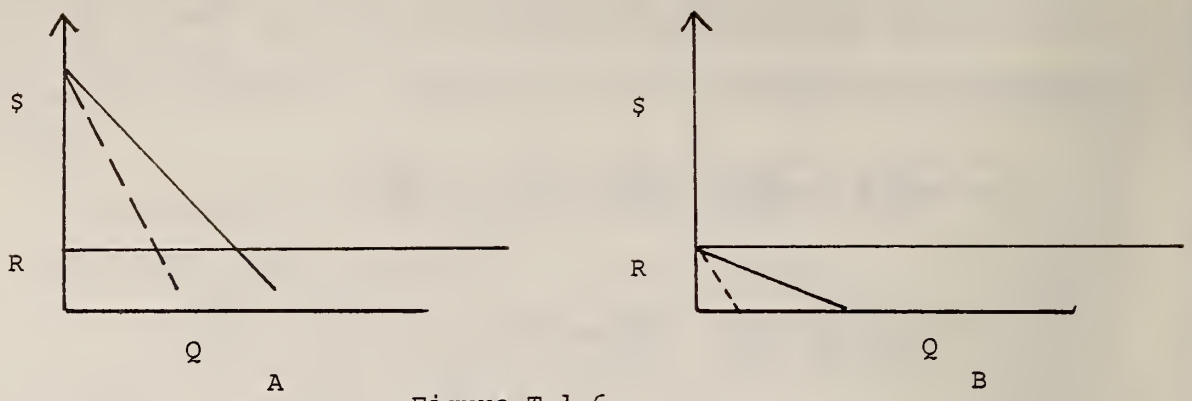


Figure T-1.6

In Figure T-1.6 suppose homes A & B start out with demand curves as shown for the same amount of service (quality) provision. It is possible that firm A can increase profits by increasing services above their current level, shifting the private demand curve still further to the right; the marginal cost of additional services per patient day will be outweighed by the increase in revenue from private patients. Firm B is currently serving all Medicaid patients. Its cost (due to increased services) of shifting its demand curve rightward are higher than for firm A, because it must overcome the private patients' reluctance to enter an all-welfare home.

These interactions imply that we are likely to see high quality services provided in facilities with a high proportion of private patients, and low quality services in facilities with a high proportion of Medicaid patients; and few facilities with proportions of private and Medicaid patients near the market proportions, because of the forces that tip facilities toward the private or the public market.

If the Medicaid rate setting mechanism decides to "pay for quality," i.e., reimburse facilities for their provision of S to patients, it is <u>not</u> clear that quality will rise in all facilities or that more Medicaid patients will be accepted in high quality homes. It may turn out that high service homes will simply make more money on their Medicaid patients, while many facilities who have no hope of getting into the private market will still not find it worthwhile to add services--unless the state pays 100% of additional cost, in which case they will be indifferent with respect to their level of S, as long as they have 100% Medicaid patients.

These issues deserve further consideration.

T-1.4.2 Optimization in the Non-Profit Firms

In the same vein, the implications of the optimization process for nonprofit facilities can be explored, given assumptions about their utility function.

Assume that firms maximize

$$U = U\left[q\,(S, x), Q_m, Q_p\right]$$

This allows private and public patient days to have different utility values to the operator in and of themselves. If the marginal utility of serving a private patient is everywhere the same as the utility of a public patient, $\frac{\partial U}{\partial Q_m} = \frac{\partial U}{\partial Q_p}$.

Subject to net revenue constraint with net revenue equal to zero (Alternatively, a deficit, $D < 0$ can be allowed.)

- a production function
- a private demand function

$$P * Q_p + R * Q_m - wL - rK = 0$$

$$Q_p = Q_p\,(P, q, Q_m)$$

$$Q_p + Q_m = \phi\,(K, L, S, A)$$

The Lagrangian is

$$L^* = U\left[q\,(S, X), Q_m, Q_p\right] - \lambda_1\,(P*Q_p + R*Q_m - wL - rK) - \lambda_2\left[\phi(K, L, S, A) - Q_p - Q_m\right] \qquad (1)$$

$$\frac{\partial L^*}{\partial P} = \frac{\partial U}{\partial Q_p}\cdot\frac{\partial Q_p}{\partial P} - \lambda_1\left(Q_p + \frac{\partial Q_p}{\partial P}\cdot P\right) + \lambda_2\,\frac{\partial Q_p}{\partial P} = 0 \qquad (2)$$

$$\frac{\partial L^*}{\partial L} = \lambda_1 w - \lambda_2\,\frac{\partial \phi}{\partial L} = 0 \qquad (3)$$

$$\frac{\partial L^*}{\partial Q_m} = \frac{\partial U}{\partial Q_m} = \frac{\partial U}{\partial Q_p} \frac{\partial Q_p}{\partial Q_m} - \lambda_1 \left(P \frac{\partial Q_p}{\partial Q_m} + R\right)$$

$$+ \lambda_2 \frac{\partial Q_p}{\partial Q_m} + \lambda_2 = 0 \qquad (4)$$

$$\frac{\partial L^*}{\partial S} = \frac{\partial U}{\partial q} \frac{\partial q}{\partial S} + \frac{\partial U}{\partial Q_p} \frac{\partial Q_p}{\partial q} \frac{\partial q}{\partial S} - \lambda_1 P \frac{\partial Q_p}{\partial q} \frac{\partial q}{\partial s}$$

$$- \lambda_2 \frac{\partial \phi}{\partial S} + \lambda_2 \frac{\partial Q_p}{\partial q} \frac{\partial q}{\partial S} = 0 \qquad (5)$$

Divide (2) by $\frac{\partial Q_p}{\partial P}$

$$- \frac{\partial U}{\partial Q_p} + \lambda_1 \left(P + \frac{\partial P}{\partial Q_p} * Q_p\right) = \lambda_2 \qquad (6)$$

λ_1 may be interpreted to be (-1) times the marginal utility of funds to the facility. λ_2 is thus (-1) times the sum of the utility due to an increased private day and the utility of the revenue derived from that day.

From (3)

$$w = \frac{\partial \phi}{\partial L} \cdot \frac{\lambda_2}{\lambda_1}$$

Substituting (6)

$$w = \frac{\partial \phi}{\partial L} \cdot \frac{\lambda_1 \left(P + \frac{\partial P}{\partial Q_p} \cdot Q_p\right) - \frac{\partial U}{\partial Q_p}}{\lambda_1}$$

$$\lambda_1 w = \lambda_1 \frac{\partial \phi}{\partial L} \left(\text{private marginal revenue}\right) - \frac{\partial U}{\partial Q_p}$$

The utility value of the wage, $-\lambda w$, is equal to the utility of the marginal product of labor, $-\lambda_1 \frac{\partial \phi}{\partial L}$ * private marginal revenue, plus the independent utility of an additional private day. Nonprofit facilities will thus use labor beyond the point where the wage equals the marginal revenue product of labor:

$$w = MRP_L - \frac{\partial U}{\partial Q_p} \cdot \frac{1}{\lambda_1} \qquad (7)$$

From (4), substituting in (6)

$$\frac{\partial U}{\partial Q_m} + \frac{\partial U}{\partial Q_p}\frac{\partial Q_p}{\partial Q_m} - \lambda_1 \left(P \frac{\partial Q_p}{\partial Q_m} + R\right)$$

$$- \frac{\partial U}{\partial Q_p}\frac{\partial Q_p}{\partial Q_m} + \frac{\partial Q_p}{\partial Q_m}\lambda_1 \left(P + \frac{\partial P}{\partial Q_p} * Q_p\right) - \frac{\partial U}{\partial Q_p}$$

$$+ \lambda_1 \left(P + \frac{\partial P}{\partial Q_p} * Q_p\right) = 0$$

collecting terms,

$$\frac{\partial U}{\partial Q_m} - \lambda_1 \left(R - \frac{\partial P}{\partial Q_p}\frac{\partial Q_p}{\partial Q_m} Q_p\right) = \frac{\partial U}{\partial Q_p}$$

$$- \lambda_1 \left(P + \frac{\partial P}{\partial Q_p} * Q_p\right) \qquad (8)$$

The utility of an additional public patient plus the utility derived from the additional revenue, R (net of the revenue from loss of private patients where $\frac{\partial Q_p}{\partial Q_m} \neq 0$) is equal at the optimum to the marginal utility of a private patient plus the utility of the additional revenue to be derived from private patients.

When $\frac{\partial U}{\partial Q_m} = \frac{\partial U}{\partial Q_p}$ everywhere, which would be true if the organization had no particular preference for public as opposed to private patients, or vice versa, (8) reduces to

$$R - \frac{\partial P}{\partial Q_p}\frac{\partial Q_p}{\partial Q_m} Q_p = P + \frac{\partial P}{\partial Q_p} Q_p \qquad (9)$$

The marginal revenue of a public patient, R, is equal at the optimum to the marginal revenue of a private patient, with the revenue impact of the consumption externality, shown by $\frac{\partial Q_p}{\partial Q_m} > 0$, as a wedge between them. (If $\frac{\partial Q_p}{\partial Q_m} = 0$, the patient days are produced such that marginal revenue in the private market equals the rate.) Note that this is the same in the for profit-analysis.

Substituting (6) into (5) and collecting terms,

$$\underset{(+)}{\frac{\partial U}{\partial a}} \cdot \underset{(+)}{\frac{\partial q}{\partial s}} + \underset{(-)}{\frac{\partial P}{\partial Q_p}} \cdot \underset{(+)}{\frac{\partial Q_p}{\partial q}} \frac{\partial q}{\partial s} \lambda_1 Q_p =$$

$$- \underset{(+)}{\frac{\partial U}{\partial Q_p}} \underset{(-)}{\frac{\partial \phi}{\partial s}} + \lambda_1 \frac{\partial \phi}{\partial s} \left(P + \frac{\partial P}{\partial Q_p} Q_p\right) \qquad (10)$$

The left-hand side shows the utility gain from an increase in services per patient day at the margin. This comes from the utility of quality per se and from the utility of the additional revenue that can be derived because with S higher, private patients may be charged more without a reduction in demand. The right-hand side underlines the fact that when services per patient day increase, variable inputs must be diverted from patient day production at the rate $\frac{\partial \phi}{\partial S}$; the right-hand side equals the utility of the loss of private days per se and the utility of the revenue loss.

Of course, under our objective function assumption, funds have value only as they produce additional quality or patient days.

Again, it appears that the model is consistent with outcomes in the nonprofit sector of the nursing home industry, specifically the larger size of nonprofit facilities and their greater use of labor inputs. These issues are worthy of further consideration.

T-1.5 Costs

The cost function relating Q and other characteristics to total costs (TC) and average cost emerges from the optimization process.

$$TC = \sum_i r_1 k_1 + \sum_j w_j L_j$$

The L_j are determined in the optimization process while capital is given, so that the following variables may be assumed to have an impact on operating costs OC.

$$OC = f(Q, A, q(S, X), y_1 \ldots y_n)$$

where $y_1 \ldots y_n$ are facility characteristics. Average operating costs are $\frac{OC}{Q}$.

Several researchers have attempted to relate average or total costs to nursing home characteristics and production (Q,A). Quality and case mix have not been treated adequately, although the conclusion can be drawn that case mix is related to costs. Patient mix by method of payment has been treated as an exogenous variable; as shown above, the public-private mix of a facility probably depends on its pricing policy, service production, and demand from private patients, as well as on the public rate.

Work by Skinner and Yett (1973, and undated reports) has explored the impact of patient and facility characteristics on costs. They carried out two analyses: a pilot study using observations on 46 facilities with the facility's average of total costs as the dependent variables, and a national study using patients as the unit of observation and with patients' total monthly charge as the dependent variable.

Full results are not reported for the national study. The only conclusion that can be drawn from reported results has to do with whether or not certain characteristics of patients and facilities are significantly related to total charges. The significant findings were:

1. for a debility index of a Guttman scale type, higher levels of dependency had significantly higher charges.

2. for primary diagnoses, only musculo-skeletal and mental classifications were significantly different from the other ten diagnostic classes: the first was higher in charge, the second lower.

3. regions may be ranked in expense from lowest to highest as follows: south, north central, west, and northeast.

4. for-profit facilities charged more than non-profit.

Results on interaction terms appear in an unpublished appendix not available to us.

These results are equivalent to regression analysis on dummy variables taking one characteristic at a time, and do not show impact of a characteristic holding other attributes constant. It is questionable that charges can be used as a proxy for costs, due to Medicaid regulation restricting payment for Medicaid patients.

Skinner and Yett (1973) report "preliminary" results for a cost function based on the national data. The equation estimated is

$$\log C = b_o + b, A + b_2 \log Q + b_3 u + \Sigma b_1 x_1 + \Sigma b_j y_j$$

where C = average costs
A = admissions
Q = patient days
u = occupancy rate
x_i = patient proportion in each debility class
y_j = other adjustment variables
b_i = estimated coefficients

Statistics on the full regression are not presented and the only coefficients are shown for the debility index dummies for separate equations for for-profit and non-profit homes. In these, only an intermediate debility class is shown to have a significant impact on costs, while those both above and below in the dependence scale had no significant cost impact. This result is peculiar if the scale is measuring dependence as described, since patients in the next dependency class have all the dependencies of the significant class plus some others. Moreover, the elasticities at the means of cost with respect to debility level of for-profit and non-profit homes are .02 and .04 respectively--very small indeed.

Skinner and Yett's (nd) initial exploratory study of cost data for 46 for-profit nursing homes reports results more completely but suffers from several basic flaws. The model used is

$$\log TC = b_o + b_1 A + b_2 \log Q + b_3 u + \Sigma b_1 x_1 + \Sigma b_j y_j$$

A stepwise procedure was apparently used, so that variants contain different subsets of the independent variables, R^2 was high. Patient proportions that had a significant effect on total cost included:

- private pay support (+)
- discharged to home (-)
- died (-)
- diagnoses
  - skin disorder (-)
  - metabolic disorder (-)
- debility
  - independent (-)
  - dependent in bathing and ambulation (+)
- female (+)

Facility characteristics in this particular regression were facility age (+) and number of shifts with licensed vocational nurses (+). Occupancy rate is treated by Skinner and Yett as if it were an exogenous facility characteristic, perhaps realistic if beds are given; but to then interpret the results as showing optimal size (given a mean in the sample of 83% occupancy) cannot be correct. They find minimum average cost at an admission rate of 173 per year--but this result also depends on a constant length of stay (p. 15). A realistic cost model must recognize that operators can vary both admission and length of stay--one would have thought that this was the very point of treating A and Q as separate variables in their cost functions.

Ruchlin and Levy (1972) regressed facility characteristics on average costs for a pooled cross section of Massachusetts nursing homes for years 1965 through 1969. Components of variable costs such as laundry, nursing, and housekeeping costs, are also used as dependent variables. Occupancy rate is used as an independent variable, as is percent public patients; both had significant negative effects on average cost. Scale measured in bed size was not significant in predicting total costs, but had a significant impact (both positive and negative) on various cost

components. Dummies for time periods are interpreted to show an impact of change in rate setting method, holding facility characteristics constant. The ownership variable showed significantly higher non-profit costs, and older homes were found to be more costly, as were facilities certified for Medicare. This study does not differentiate among levels of care, except for the certification variable, and does not account for patient characteristics except for payment method. Since accounting data are used, costs, especially depreciation, are sensitive to variations in accounting methods. Dummy variables for time periods do nothing to explain any change through time.

These results can be used to point up association between cost and various facility characteristics, but cannot be extrapolated to show what would happen to costs if one variable, like bedsize or percent public pay, were to change. For example, facilities with a high proportion of public pay patients tend to have lower costs; this is explicable using the model under development here, in that firms with little hope of attracting private patients do not find it worthwhile to produce "extra" S, and thus have low costs, few private patients, and many public patients. Low occupancy rates may be associated with high costs not only because of economies of scale but because facilities serving mostly private patients, and thus providing a high level of services per patient day, find that serving additional Medicaid patients to increase occupancy reduces private demand. The impact of case mix on cost cannot be delineated clearly without attention to the role of quality, or adequacy of services for the case mix. Walsh (1976) makes this point graphically:

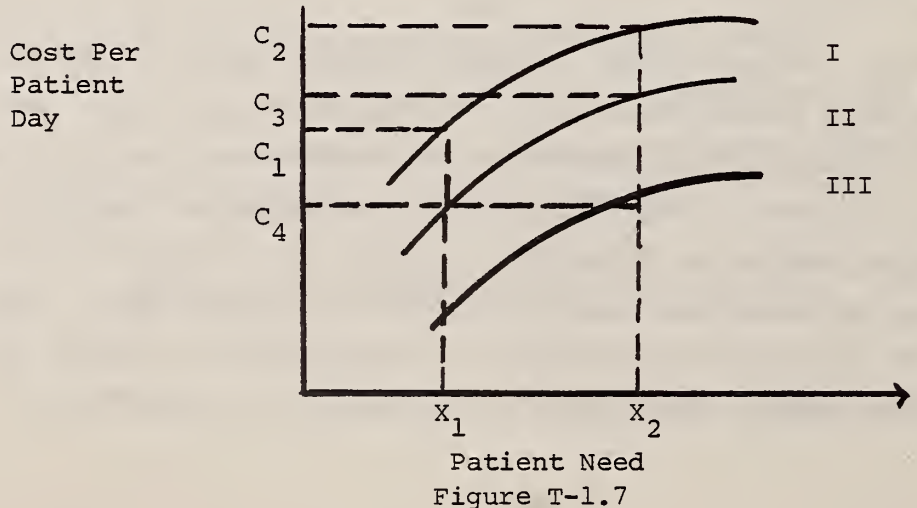


Figure T-1.7

In Figure T-1.7, facilities I, II, and III are operating at three levels of services intensity, with I with the largest amount of services per patient day. As patient need increases from X_1 to X_2 in facility I, coverage cost increases from C_1 to C_2; but case mix X_2 will only cost C_3 in facility II and C_4 in facility III. The independent impact of case mix on cost cannot be estimated from a cross section of facilities at many levels of service intensity and varying case mix without a way to hold service intensity constant.

This argument is supported by the finding of Winn, Bennett, and McCaffree (1976) that related patient condition determines the proportion of staff time received by each patient, rather than the absolute amount. This suggests a reallocation of staff time in response to case mix changes, rather than an increase in inputs (and thus cost). The missing variable in the cost functions including case mix is quality; alternatively, if a measure of services can be found, this should replace both case mix and quality, since $q = q(S, X)$, and it is the intermediate good S that actually is calling forth resources and adding to cost.

T-1.6 Summary and Conclusions

A model of firm behavior such as the model developed in Section T-1.4 is essential under-girding for the development of a prospective reimbursement system. The model put forward here captures a number of interesting aspects of firm decision-making that make it especially useful for understanding both broad issues concerning the impact of regulation on the industry and more narrow questions with respect to cost function estimation.

T-1.6.1 The Role of Quality

It is a shortcoming of previous work on nursing home costs and production that quality of care is not dealt with directly. The model above assumes that facilities do not produce quality itself, but a set of services per patient day, S; only when these are supplied to patients with a given debility mix X, can a determination be made about the quality of care produced. In cross-sectional analysis, under our model, it is an increase in S that raises cost per patient day, not an increase in X (a worsening case mix which need not call forth additional inputs) or an increase in q (which could be the result of a decline in X, with inputs constant). This implies that cost functions must hold S constant, or use the relationship $q = q(S,X)$ to find the inverse function $S = S(q,X)$ and hold constant both q and X; only in this way can cost functions be assumed to represent the relationship between cost and a homogeneous good. The interrelationship between q, S, and X also highlights that regulators are ideally seeking a floor under q through regulations that attempt to control S and X. Regulators specify appropriate public patient placement for a level of care, but this provides only partial control over X, since private patients may seek care at any certified level. Regulators also specify minimum levels of inputs, like nursing homes per patient day, which, if employed efficiently by the firm in production of S, determine a minimum level for S. Regulators do not promulgate standards for patient outcome or patient comfort, but attempt to set a standard for q using S and X.

Facilities have an incentive to produce more services per patient day than regulators require for their level of care if private patients, as postulated above, value quality, where $q = q(S,X)$. These incentives are

probably stronger for facilities with an initial advantage in attracting private patients, due to location or physical plan attributes. In like manner, some facilities have an incentive to keep the difficulty of their case mix low. This way of looking at q, S, and X will lead to testable hypotheses about S and X for facilities of various types, and suggests that regulators interested in purchasing a certain quality level for public patients cannot rely on input standards and partial control of case mix to meet this goal.

T-1.6.2 Medicaid Patients as a Deterrent to Private Patients

People appear willing to pay throughout their lives to assure that their neighbors are of similar social status. It would be surprising if such behavior ceased at the nursing home door, and it does appear that some nursing homes, like carefully zoned suburbs, do cater to a relatively well-off portion of consumers. It is hypothesized that the number (or percent) of Medicaid patients served by a facility has a negative effect on private demand. This must be tested using national and state data, but such a "tipping point" notion may well help to explain why there are empty beds in nursing homes serving a high proportion of private patients, while homes serving mostly public patients are fully occupied.

T-1.6.3 Case Mix

It is implicit in the model that a more needy case mix in itself does not call forth more services. In addition, facility operators appear to have some control over case mix. Therefore, debility level of patients may be used as an explanatory variable in a cost function only with extreme caution: if facilities serving a high proportion of relatively "well" private patients by providing high quality services, including amenities, at relatively high cost per patient day are able to screen out the most disabled public patients, it will appear that less disabled patients cost more per patient day.

A rate that is higher for more disabled patients is an incentive for keeping patients dependent, while providing a reward for accepting more disabled Medicaid patients does not in itself encourage operators to supply the services that these patients would need for high quality care.

T-1.6.4 The Cost of Care

Firms that produce more services per patient day are likely to have higher costs per patient day. As discussed above, more services probably also means more private patients. Thus several variables will probably be intercorrelated: percent private pay, occupancy rate (inversely correlated), and cost per patient day. A cost function that treats occupancy rate and percent private pay as exogenous variables will miss this interrelationship: the facilities are not high cost because they have many private patients, rather they are able to attract private patients because they are spending to provide a high level of services per patient day. If a reimbursement system increases reimbursement for services, this may increase the facility's production of these services--and increase its proportion of private patients.

T-1.6.5 The Private Price for Nursing Home Care

The model postulated above implies that nursing homes produce differentiated products so that each faces a downward sloping demand curve and can choose a price-quantity combination in its particular market. It is not assumed that all firms face the same demand curve, so that characteristics of each particular facility affect the position of its private demand curve. The intersection of the marginal revenue curve and the Medicaid rate determine the profit maximizing quantity, and thus the price, of private care--allowing for the externality in consumption discussed above. The relationship between private price and marginal cost is tenuous at best, since firms differ in their market power and the relevant marginal cost is not the cost of production but the opportunity cost of a private day, namely the revenue from a public day. Thus while private price will be an important variable to explore, it is not a valid input into cost-based reimbursement (compare Deane and Skinner, 1977).

T-1.6.6 The Endogenous Nature of the Medicaid Rate

Where Medicaid rates are based on the per diem average costs of individual facilities, the public rate can be influenced by firm behavior. More services per patient day implies higher quality (case mix constant) and thus greater private demand; it also implies higher costs, and thus a higher public rate. If costs are higher because a facility is operating at less than optimal capacity utilization, or has a non-optimal number of beds, a cost-based rate may inappropriately recognize these costs. The model of nursing home behavior points up these pitfalls for a rate setting system.

T-1.6.7 Capital Costs

It has been argued above that consideration of the method of reimbursing capital costs can legitimately be severed from the development of a method for reimbursing operating costs. This is because a new entrant into the nursing home industry will choose a profit maximizing bed size and the like based on expectations about total return to capital (for example, appreciation of real estate, reimbursement of depreciation and property taxes, and opportunities for investment of expected cash flow). Once the facility is in place, however, the operator will maximize the difference between operating revenue and operating cost, and cannot change capital investment from year to year in response to regulation changes. However, both capital and routine cost reimbursement determine the profitability and hence the viability of firms. Current methods of reimbursement of capital expense, with their rewards to debt financing and rapid depreciation, may favor a higher ratio of capital to variable inputs than would otherwise be the case. This may mean that the profit maximizing size for a new facility under current reimbursement rules is larger than the optimum efficient size (minimum average variable cost). The variation in capital reimbursement rules across states may well provide enough variation in scale for each vintage of capital to enhance econometric results. However, if an optimum scale is found, it is probable that reimbursement methods alone cannot push facilities toward it, but that determination of need regulation must be used in addition to encourage efficient sized facilities.

T-1.6.8 The Nursing Home Industry Under Regulation

Regulation shapes the behavior of individual firms and in this way influences the configuration of the entire industry. The rate of entry and the rate of survival of facilities in different size classes, ownership groups, and geographic locations will be affected by regulators' decisions in the future as they have been in the past. A review of the evolution of the nursing home industry (Chapter 2) showed a cottage industry which all too suddenly became big business with high profits. Under increasingly intense regulation came a growth in the average size of firms, the emergence of nursing home chain ownership and nursing home management companies, and public demands for higher quality

which occasionally have been translated into favorable treatment of nonprofit facilities, which are felt to provide high quality care. Data are not available as yet to evaluate market segmentation of facilities, but it appears that some firms position their product to serve the private market while others are strictly Medicaid. Regulators must be aware that these dimensions of industry behavior may change in response to rate regulations directed toward some other purpose entirely.

For example, it appears that nonprofit facilities, because they value quality, will use more variable inputs and produce more services per patient day than for-profit facilities facing the same facility demand schedule and the same rate. If rates are designed to reimburse in full for services provided to patients (while perhaps limiting other costs, like capital costs) nonprofit facilities will be further encouraged to "overproduce" services, maintaining their cost differential above for-profit homes. Rather than backing into this outcome, an explicit determination should be made about whether these extra services per patient day are worth their cost, and if so, whether for-profit facilities also should be rewarded for providing them. The role of private philanthropy in supplying high quality nursing home care should be evaluated so that public policy neither eliminates this source of resources nor in effect gives private philanthropic organizations control of the allocation of public resources.

Secondly, it is important to understand, both theoretically and empirically, more about the role of private-public market segmentation. It is possible that rate regulations are bringing about a two-class system of nursing home care, made up of homes serving mainly private patients and homes serving mainly public patients. Alternatively, the private patients can be a force to "police" the value and quality of nursing home care. Even more attractive, with advocates or ombudsmen and more alternative types of care, both private and public patients could become truly sovereign consumers. In such a case, the industry would find that attributes of care valued by patients and their families were rewarded.

In any event, the nursing home industry ten years from now may well look as different from that of today as it is different from nursing home industry of ten years ago. Government regulation, particularly rate regulation, is sure to play a large role in determining these changes.

TECHNICAL PAPER #2

Technical Paper #2

REIMBURSEMENT SYSTEMS IN OPERATIONS*

T-2.0 Introduction

The purpose of this discussion is to review existing and recently proposed (1978) Medicaid nursing home reimbursement systems in operation across the nation. Since the focus of this report is an analytical review of the determinants of nursing home costs and how a reimbursement system might best take these cost factors into account, an indepth analysis of current plans is not appropriate. Rather, the discussions in Sections T-2.1 and T-2.2 review the major variations of different reimbursement systems.[1] Section T-2.3 contains an extensive critique of some of the more common grouping procedures used to classify homes for cost comparison purposes.

At the outset it should be noted that it is not possible to develop a meaningful typology of reimbursement systems. The reason is that the existing systems vary along a broad range of dimensions. For example, two states which both have prospective reimbursement systems may differ widely in their methods of determining the prospective rate as well as their methods for grouping facilities and imposing cost ceilings. Consider the cases of Texas, West Virginia, and Rhode Island. In all three states, rates are set on a prospective basis, i.e., in advance of the year for which the rate is to apply. In Texas, facilities are reimbursed uniformly, depending on which classification group they fall into (in terms of level of care, size of facility, and geographic area. In contrast, Rhode Island and West Virginia reimburse prospectively on the basis of historical costs of each individual facility. But West Virginia and Rhode Island themselves differ in terms of

---

*by Christine Bishop

[1]The source for this discussion of nursing home reimbrusement systems is forty descriptions of current proposed state reimbursement systems supplied by the Medicaid Bureau of the Health Care Financing Administration (1978). Descriptions of the remaining ten states were not available. The other major reference sources on historical nursing home reimbursement systems are Applied Management Sciences (1977), the Urban Institute (1977), and Social and Rehabilitative Services (1973) which are discussed in Technical Paper #4.

the maximum ceilings they put on costs. West Virginia develops these ceilings on a "peer group" basis (in terms of classification level, size, and geographic location); Rhode Island develops these ceilings by considering all the homes in the state. Another difference between the two states is that West Virginia has an efficiency incentive (which shares with the facility the difference between the provider's cost and the cost ceiling) while Rhode Island provides for no efficiency incentives. Similar wide variations in reimbursement practices pervade the other states' reimbursement systems.

There are two other alternatives for organizing this review of existing reimbursement systems. The first is to discuss in detail a few states which encompassed the extremes of possible reimbursement system. But in practice, this alternative, too, deteriorates into a non-generalizable set of several case studies. The other alternative, which is followed below, is to discuss each of the components which are involved in a reimbursement system. Instead of using the state as the unit of analysis, the remainder of this Technical Paper considers each of the constituent parts of a reimbursement system as the unit of analysis. In this way, a more complete portray of existing reimbursement systems is possible.

T-2.1 The Parts of a Reimbursement System

The distinctions among reimbursement systems can be broken into several constituent parts: (1) the parameters of the system; (2) the administrative mechanisms; (3) the discriminating variables; and (4) the audit process. These terms are defined below and are operationally discussed in more detail in Section T-2.2 (with the exception of the audit process, for which no data were available).

Parameters

The parameters of a reimbursement system are the normative standards that define how the price is to be established. In the context of nursing homes, the parameters or norms involve the procedures for defining "reasonable" cost. These norms may be directly or indirectly stated. The chief form of direct parameters are the input requirements that are placed on homes: there must be so many nurses, buildings must meet specified life and safety codes, administrators must have specified training. These direct parameters encompass the range of licensure and certification standards required for

Medicare and Medicaid reimbursement. It should be noted that these direct parameters typically take the form of input requirements and can not be translated directly into measures of output performance. In essence, these direct parameters specify a range of expenditures which are "related" to patient care and are reimbursable.

Indirect parameters involve procedures for defining the "reasonableness" of the particular expenditures which meet the direct norms. In essence, the indirect parameters state whether reasonableness is to be considered on a facility dependent basis (i.e., considering the costs of only the home in question) or a facility independent basis (considering the costs of the home in the context of other homes in a specified group), and the level of independence to be allowed, i.e., the size and composition of the group that a particular home is to be compared with. For example, there is discussion of whether the full costs of all related expenditures of a particular home are to be included, or based on the mean (or median or some other parameter of the cost distribution, e.g., 80th percentile) of homes in a particular category of characteristics (a uniform class rate), or based on some recognition of individual home differences in a particular group (a ceiling). For the most part, the direct and indirect norms are both a function of the political environment. If scandals and fraud are making political headlines, the norms are apt to be strengthened, as was the case in New York after the Moreland Commission report. The parameters also appear to be a function of how many homes are needed in a state and on the state budget.

Administrative Mechanism

The administrative mechanisms define the process of setting the rate and involve the nature of discretion allowed in the system. A key dimension of the administrative mechanism is the length of time for which the administrative price is set. At one extreme is a retrospective system that may recognize rather immediate changes in cost.

This contrasts with prospective systems that set prices for a period of time in advance, say one year. Another dimension is the actual process under which the rate is based; it may be based on a home by home review of submitted cost forms or through budget reviews, or through statistical analyses of data from groups of homes. Other aspects relevant to the administrative mechanism involves the nature of the data collection forms and the composition of individuals who make up the rate setting agency.

Discriminating Variables

A third aspect of a reimbursement system is the discriminating variables that are used to distinguish homes. The discriminating variables involve equity decisions about what is a fair way to recognize relevant differences among homes in the reimbursement process. Once these differences are established, they become incorporated in rate setting activity according to the parameters and administrative mechanisms discussed above. Typical discriminating variables include the size of the home, its occupancy rate, input prices, services, and other characteristics of the home.

It is these discriminating variables which are used by the rate regulators to establish ceilings and caps on reimbursement. As a consequence, the discriminating variables are the key ingredients or characteristics of the reimbursement system which lead to the resultant market outcomes. Because of their importance in establishing maximum rates (which in turn became the effective rates), most of the state-specific discussion below focuses on the range of discriminating variables actually used by the states. Indeed, the major focus of the empirical work in this study has been to examine these discriminating variables in terms of the direction and the magnitude of their expected effects. Section T-2.4 expands on this discussion to suggest which ones are useful and appropriate and which appear to be inappropriate and lead to inefficiencies or other undesirable consequences.

Audit Process

Every reimbursement system includes a variety of mechanisms which audit the actions of the providers as a measure of control. The audit function is concerned with how well nursing homes are performing. Thus, the audit processes (and their resultant rules and regulations) are designed to insure that licensure and certification standards are met as well as to insure proper financial practices and the elimination of peculation and other forms of scandal. Inspectors, accountants, and other analysts check up on the nursing homes in efforts to avoid financial fraud and abuse as well as patient care abuse.

An important thing to note is that the audit rules can be subjectively interpreted by the auditors. It appears that the rules are differentially applied to homes, at least in part, as a way of subjectively correcting conditions which may (or may not) meet the letter of the law but are in

fact distortions of what is actually occurring. For example, Cohen (1978) relates that in Maryland it is a widespread practice among inspectors to overlook infractions of the rules in homes which in fact are producing good quality care while making sure that homes which are producing extraordinarily poor quality somehow fail the inspection, even if they meet the rules on a superficial basis. Similarly, an additional element affecting subjective interpretation is the nature of excess demand in an area. The larger the waiting lines, the more difficult it is to close homes or stop the placement of public patients. While the home may be inadequate, an apparently common attitude among inspectors is that a bed in a low quality home is better than no bed at all.

T-2.2 Reimbursement in Practice: Parameters and Administrative Process

This section illustrates some of the existing parameters and administrative processes currently in use or recently proposed across the states.

Information on the direct parameters of what inputs are reimbursable (such as dietary, nursing, and life safety requirements) was not available. However, there was information on the type of indirect parameters used by the states to establish what is reasonable. A useful typology is whether they rely on the full costs of individual facility, or on uniform rates set for classes of facilities, or on a mixed strategy of full cost reimbursement for facilities below a ceiling and rates set at the ceiling (a uniform class rate) for facilities with costs above the ceiling. These terms are developed more formally below.

Full cost reimbursement refers to payment rates which follow the Medicare payment statute that payment be on a "reasonable cost-related basis." This has been interpreted to mean that payment may not exceed reasonable changes consistent with efficiency, economy, and quality of care. (PL 92-603)

Uniform class rate systems use a measure of the central tendency of costs as the single rate for facilities within a given group. For example, Mississippi uses the mean, while Tennessee uses the median.

Ceilings allow for individual variation that is not recognized by placing facilities in groups. Ceilings typically are based on the mean or median for a group. Then, a ceiling is added on. Typical ceilings are one standard deviation (New Mexico), a percentage of mean (Minnesota: 25 percent), or a percentile (Wisconsin: 55th percentile).

With regard to administrative process, the key issue of concern here is with regard to timing--whether the system is prospective (where a facility's rate is set for the next year) or retrospective (relying on current costs). While information on the nature of the timing was available--see Table T-2.1--there was no information on data collection instruments or rate review staff. While the process used to define the rate is of interest, it can be summarized quite briefly. The process for states using principles of full cost reimbursement is a home by home review of the cost reports. With the exception of Washington which has developed a formula based on multivariate cost analyses similar to those discussed in Chapter 3, all other states (which fall into the uniform or ceiling rate dimensions) conduct tabular analyses of homes in each of the groups.

Table T-2.1 displays forty states for which data were available in terms of these dimensions.

A number of states elected to follow Medicare full cost reimbursement principles. Full cost retrospective reimbursement was chosen for Skilled Nursing Facilities in two states, while five others choose to reimburse all their homes in this fashion. Two Medicaid programs, Montana and the District of Columbia, used Medicare principles to set the base for individual facility prospective rate. This cost base for each individual home then is projected into the rate year using an inflation factor. Since over time such a method eventually will recognize individual facility costs, it is equivalent to retrospective full cost reimbursement with a lag.

Four states set uniform class rates for facilities; in two other plans (Montana and Texas) uniform class rates were used for some components of costs. No state uses a retrospective uniform class rate.

The majority of the methods proposed relied on ceilings to screen costs for reasonableness. Ceilings are in general based on historical costs for the group, fulfilling the requirement that rate setting be "cost related." In ten states, they are applied to historical cost for retrospective reimbursement; facilities below the group ceiling receive full cost reimbursement while those above it receive only the ceiling. In twenty-one states, a historical ceiling screens past costs, which then is projected into the rate year to form a prospective rate for each facility; homes with costs above the ceiling are assigned a prospective rate equal to the ceiling times the inflation factor.

Table T-2.1

	Retrospective	Prospective
Full Cost (Medicare)	Alaska Hawaii Idaho* Kentucky (SNF) New Hampshire (SNF) Utah* Vermont*	Montana I District of Columbia**
Uniform Class Rate		Illinois Mississippi Montana+ Oklahoma Texas Wisconsin+
Ceiling	Maryland I Massachusetts Michigan I Montana+ Oregon Pennsylvania I Tennessee (SNF) Utah Virginia Washington	Alabama Colorado I Delaware Florida Georgia I Indiana Kansas Kentucky I (ICF) Minnesota New Hampshire I (ICF) New Jersey New Mexico New York North Dakota Ohio Rhode Island South Dakota Tennessee (ICF) West Virginia I Wisconsin+ Wyoming I

KEY:

I = incentive plan
+ = some components only
* = full cost with ceilings on certain administrative components
** = administrative cost limited to 15% of total

A further aspect of current and proposed reimbursement systems is the use of incentives to facilities with costs below the ceiling. For three of the ten retrospective ceiling methods and six of the twenty-one prospective ceiling methods, incentives are included for keeping cost below the ceiling. The incentive involves sharing some portion of the difference between the incurred costs and the ceiling rate with the home (if costs are below the ceiling). These states are indicated with an I in Table T-2.1.

In sum, major prototype reimbursement systems are full cost reimbursement methods, uniform class rate methods, and ceiling methods. Both the uniform class rate and ceiling methods face the problem of determining a particular rate which either becomes the rate for all facilities in the class, or for all facilities in the class that exceed the rate. Hence, assigning facilities to classes, and determining the amount by which the rate or the ceiling will deviate from some measure of central tendency for that class are the discriminating variable issues that must be settled. The incentives to keep costs below the ceiling will also affect the outcome of ceiling methods with respect to cost containment. The next section (T-2.3) presents an analysis of variables that are typically used as grouping criteria. These include the cost variation due to scale of production, utilization of capacity, services, case mix, input prices, and type of patients served. Section T-2.4 discusses current rate setting methods for treating cost related variables in light of the empirical analyses conducted by this study (in Chapter 3).

T-2.3 Reimbursement in Practice: Discriminating Variables

A variety of discriminating variables are used by the states to distinguish among nursing homes for purposes of establishing reimbursement rates. In establishing these discriminating variables, rate setters are implicitly and explicitly stating which type of cost variations are acceptable and which are not, that is, what product dimensions will and will not be paid for. In a typical group ceiling process, the data used to set the ceiling are drawn from the same group to which the ceiling is to be applied. This is apparently a way to fulfill the requirement that rates be cost-related: ceilings used to screen costs are not arbitrary, and may not be based on state budgets, but are based on actual costs of the facilities to be screened.

In no state is the rate tied directly to what patients actually receive in a facility, i.e., to output measures. The closest that states come to tying rates to patient care is in considering inputs of the home (which were discussed above, as minimum parameter input standards). Instead, these implicit product dimensions refer to facility characteristics (ownership, size, location, . . .) and general levels of inputs provided. However, ceilings tend to be more generous for components of cost that are believed to be related to quality of care. Thus, under some systems a facility might reach relatively low limits early on for capital costs but find most nursing and personal care costs reimbursed in full. Such a policy does not directly encourage high quality care, since no incentive is paid for such care; but it does support higher service intensity care which may or may not be related to high quality care. The following paragraphs discuss typical ways found in the states to use these discriminatory variables. These then can be assessed against the analysis of cost data.

Scale

The size of facilities or scale, measured in number of beds, is used to group facilities for the application of cost ceilings for both prospective and retrospective rate setting either for total cost or various components of cost.

Apparently, states which use bed size as a grouping variable for a given cost component believe that average costs vary with scale for that component. For example, Minnesota, Kentucky, Indiana, Utah, Massachusetts, and New Hampshire impose different ceilings on total allowable management costs or an owner's compensation based on facility size. The allowance per available bed day declines as the number of beds rises, implying that economies of scale are assumed. A similar phenomenon occurs for vehicle expense for Rhode Island. Massachusetts groups facilities for determination of nursing cost ceilings according to whether they have 40 beds or fewer, or have more than 40 beds. One must infer that the mean and/or standard deviation of nursing cost per patient day are expected to differ between these two groups, but the regulations do not specify which group will have the larger ceiling. Other states (including New York, South Dakota, Rhode Island, Texas, West Virginia, Alabama, Georgia, Illinois, and Kansas) set different overall

ceilings for different size groups, again with the implication that these are expected to be different. Finally, it can be assumed that the one state (Washington) proposing to set rates using regression of average cost on all cost "factors" with a significant level better than 20 percent on average cost could conceivably include beds in its formula if that variable met the significance criterion.

Utilization of Capacity

Occupancy rate is another variable believed by many to affect average cost in nursing homes. The way it is used in current rate setting mechanisms also allows inferences about how state regulators view the impact of utilization on cost.

Implicit in most systems is the idea that production is inefficient below a certain occupancy rate, i.e., more costly than necessary and that these "extra" costs, unlike costs that are due to firm size that is larger or smaller than the least cost level, should not be paid by the public payment program. The usual way for the occupancy floor to act is that total cost, or a specific component of total cost, is divided by actual patient days or the patient days that would have been produced had the home been operating at the required occupancy level, whichever is larger. Because of this discounting procedure, a facility operating at less than the occupancy floor cannot possibly have its actual average total cost covered by reimbursement. This seems highly appropriate in the case of fixed costs: low occupancy means a higher actual average cost than would occur at a more optimal level of utilization, and the implicit decision is that the public program should not pay these higher costs. For example, North Dakota, applies an occupancy floor to components of operating costs which a priori are expected to vary little with volume, namely, administrative costs, plant operation costs, and property costs. New Jersey applies a 95 percent floor to costs that can be seen as relatively fixed in the short run, for example utility cost. Pennsylvania, Minnesota, Utah, Georgia, Idaho, the District of Columbia, and Indiana also apply an occupancy floor to costs that are fixed given facility size.

Most other states apply the occupancy floor to total average cost, with different implications. For costs that truly vary with output, the occupancy floor implies a penalty to low occupancy not based on the average

cost at different occupancy rates; a facility operating at 80 percent occupancy instead of an occupancy floor of 90 percent is reimbursed only 8/9 of the actual additional variable cost of caring for a patient. (This fraction approaches 1 as occupancy approaches the floor, so the penalty falls.)

In addition to using occupancy floor information for analyses of the cost of a single home, they are sometimes used, as in Illinois, for comparisons across homes. In essence, the average operating costs for low occupancy facilities first are discounted by their occupancy deficit before being added in with the costs of other homes, to determine average costs for the group as the basis for determining a ceiling. One might draw the inference that state regulators believe these low occupancy facilities have abnormally high operating costs, although operating cost is assumed by most people to be variable with volume. The inclusion of these discounted costs from low occupancy homes serves to lower the mean for all facilities and thus affects the ceiling calculation.

Some state systems also treat low occupancy homes as abnormal, but exclude them from computations for this reason instead of trying to discount their costs. An example is Mississippi which includes only facilities of greater than 85 percent occupancy in its computation of the average cost increase for the industry.

Services

Services provided like nursing and personal care hours are a means of input differentiation of the product of nursing homes. As in other industries, the costs of these different "products" may be expected to differ. The differences among the products provided by various nursing homes are handled mainly by grouping facilities by broad types of care provided. Different rates for SNF and ICF care are mandated by federal policy, and some states also distinguish facilities staffed to care for mentally retarded (e.g., Pennsylvania), and/or for patients with psychiatric diagnoses (e.g., New Hampshire), and facilities providing hospital-based SNF care (e.g., Minnesota). Several states (e.g., Kentucky and Tennessee) have completely different methods of setting SNF and ICF rates[1] but most states compute and apply similar types of ceilings to homes classified into the two (or more) classes of care.

---

[1]These states apply a prospective ceiling to ICF homes while allowing full cost (Medicare) rates for SNFs.

There are two basic approaches to arriving at group ceilings for the different classes of care, one relying on group data and the other by basing ceilings on the expected inputs. Arriving at ceilings by finding mean costs for each level implies that the product currently produced by the facilities is both appropriate within levels and different across levels. A variation on this approach is to find a ceiling based on the average for one group and specify that the ceiling for the other group is a certain percentage of the first ceiling. This implies that the prices should have a fixed relationship to each other, and that data on costs for one group can be used to determine prices for the other.

The second general method for service discrimination among homes takes expected inputs for each level (for example, 2.5 nursing hours per patient-day) and multiplies these by unit costs of inputs. Implicit in this rate setting method is the idea that regulators know what the product is that they wish to purchase for public patients, and have calculated how much it should cost. This may or may not be coupled with investigation of what inputs the provider actually purchases or how much he actually pays for them. For example, in Utah the nursing input hours are specified and the input price is computed for each facility based on its own price of nursing services. Wisconsin's system is a good example of a method of paying for variation caused by differences in service intensity. Like many other states, ceilings are placed on nursing expenditure by types of care provided based on minimum hours of nursing required by licensure regulation times a mean nursing wage (based in this case on the providers' actual average wage, with a maximum for this wage rate set by state civil service scales and an assumed staff mix of three licensed nurses to seven nurse aides). If more care is provided, it is reimbursed, but only up to limits, specified in hours. Thus, there are ceilings on both the hours and the average wage per hour.[1]

---

[1] An interesting variation, but one not found in practice, might be to take Wisconsin's distinction between minimum and supplemental allowances and make it more meaningful, for example, by paying all facilities for minimum input nursing hours at a market-determined wage rate, and sharing with the facility the cost of increased hours: this would encourage provision of higher intensity services, but would not, as is currently the case, pay in full for any supplemental intensity the facility chooses to buy for its patients, up to the ceiling.

Case Mix

It is assumed in case mix oriented discriminants of nursing home costs that more disabled patients receive more services and that therefore the cost of care to these patients must be higher. All the state systems deal with case mix issues through paying different rates for two levels of service intensity, ICF and SNF. The methods implicitly rely on placement mechanisms and utilization review to assure that the patients are appropriately placed and are not receiving either more or less service intensity than is deemed appropriate by public officials. Only one state method breaks this service-patient categorization down beyond two levels: Wisconsin's plan indicates a willingness to pay a supplement for patients participating in rehabilitation programs. (There is no mention of a procedure for certifying that rehabilitation is appropriate for these patients.)

Current rate setting methods have not been oriented to finding the cost of caring for patients of a given disability level. Only one proposed plan reviewed for this analysis indicated a supplemental payment for a specific type of patient regardless of the services delivered: Wisconsin proposed a supplement for patients who exhibit disturbed behavior. However, implicit in several plans is the idea that homes supplying care mainly to private patients are supplying the "wrong" kind of care (presumably, too much care) to public patients. That is to say that many states exclude homes which concentrate on private patient care from the computation of what is reasonable cost.[1] New Jersey excludes from its sample homes with 20 percent or fewer Medicaid patients. Minnesota and Pennsylvania also have different methods of determining rates for homes with very few Medicaid patients. For example, Pennsylvania prospectively negotiates a rate with homes with a monthly average of ten or fewer Medicaid patients.

Input Prices/Location

The logic behind allowing for input price variation is to allow access to patients living in high cost areas. Without compensation for their higher input costs, nursing homes in these areas argue that they

---

[1] Such provisions also may be seen as lowering administrative cost: it seems unnecessary to go through the entire rate setting procedure to set a price for a very small number of days, and negotiation, or an optional flat rate, may be performed by both the facility and the public payment program.

will lose money and ultimately have to leave the industry. A number of state plans adjust for input prices explicitly, while others use regional groupings to assure that facilities in high cost areas were not penalized and those located in low cost areas were not rewarded.

One method of arriving at nursing cost ceilings by level of care adjusts for input price variation by using the facility's own average wage rate to determine its ceiling. Several methods employ actual market wages explicitly. For example, New Jersey will apply an "equalized compensation ratio" to three market area groups, urban, suburban, and rural. This implies that the labor component of cost will be adjusted by a wage index before it is screened using a state-wide ceiling. Another way of doing this would be to apply market wages by region to nursing homes to set the nursing cost ceilings; these would then differ systematically by market area.

A related method groups counties by a wage indicator. Georgia uses three groupings, based on average weekly earnings statistics (the type of workers is not specified) but should be held constant, e.g., manufacturing, service, etc., if a reliable indicator of factor price is desired). Alabama classifies facilities into three groups based on average yearly earnings in the health sector.

A method relying on actual data from regional groups of facilities makes the assumption that any factor price differences will be reflected in actual facility costs, so that regulators need not develop the actual weights for a price index. For example, New York State screens costs using the means for four regional groups; the implicit assumption is that costs in New York City vary from costs in rural upstate New York because of factor price differences which should be recognized by the reimbursement method. A less aggregate locational grouping scheme to allow for price variation is used by Pennsylvania. Pennsylvania classifies counties into high and low cost groups based on whether county averages for nursing home costs are high or low.

Ownership

A typical argument heard from both non-profit nursing home operators and some rate regulators is that non-profit homes supply a product different than that produced by for-profit nursing homes. That is, non-profit homes, it is claimed, produce care of a higher level of quality and/or treat sicker, (presumably costlier) patients. Hence, as the argument runs, non-profit

homes should receive differential treatment. In practice, discriminating among homes along this dimension means that profit and non-profit homes are grouped differently for analyses of their costs, as in Michigan, Kentucky, and Delaware.

T-2.4 A Critique of Existing Grouping Procedures

The paragraphs below elaborate on the argument made in Section 4.2.3.1 of the report that existing methods for determining maximum caps on reimbursement rates are concerned with equity rather than efficiency issues and are an ad hoc collection of often incorrect and ill conceived impressions about what causes variation in nursing home costs.

Before discussing specific grouping criteria used to stratify homes for cost comparisons to determine what is a reasonable maximum cap, it is useful to point out several generic issues.

Methods of comparing costs across homes rely on summary statistics of the actual cost distribution--like the standard deviation or percentile--to isolate outlines. For example, the maximum reasonable cost may be defined as the average mean cost of homes in a particular group. These procedures imply that a proportion of intra-group cost variation always is acceptable to the rate setter and that a portion of it (due to costs of facilities above the ceiling) is never acceptable. If facilities with unacceptably high costs move toward the mean (or to just below the ceiling), the distribution will change. It is implicit in a percentile method that the same proportion of homes are deemed to have unacceptable or unreasonable costs year after year, even if all bunch as close to the median as they can. The standard deviation method reflects the same underlying assumption if costs are normally distributed, since the mean plus the standard deviation will occur at about the 84th percentile in a normal distribution. It should be noted that a perverse result is possible for the standard deviation method, since large deviations below the mean can increase the standard deviation and hence the ceiling. In this way, a ceiling method penalizes high cost facilities (ceteris paribus) without rewarding low cost facilities, and thus may have an inflationary effect insofar as the ceiling is located above what minimum cost would otherwise be. If rate setters can state what rate they are willing to pay for care in a facility with specific characteristics, it seems reasonable that they should not pay that amount plus a ceiling factor (like 10 percent) to facilities with higher costs, and at the same

time fail to reward homes that can provide the specified care for less. It would appear that the only way to reconcile the reasoning behind ceilings is to see them as a "fudge factor" which allows for imperfect information on the part of rate setters and reduces claims of "arbitrariness" and the need for judications.[1]

A second point is that any payment method that is facility cost-based reinforces the status quo with respect to cost and product type provided. Because the rates are based on historical costs, many facilities will have their costs covered if they continue to operate in the same way; their profits will not be increased if they produce the same output at lower cost, or change to a more desired product (e.g., care for more disabled patients). In the long run there is little impetus for care to be provided by the most efficient facilities, those who are able to provide care at least cost per unit of output. As evidence of the support of rate setting for the status quo, it appears that some grouping variables have more of a relationship to actual cost in states where they are used as grouping criteria than in states where they are not used as grouping criteria. For example, the cost analyses in Chapter 3 found that non-profit facilities are relatively less costly in Massachusetts, where a single ceiling applies to the profits and non-profits, than they were in New York at a time when separate group ceilings were applied.

In a similar vein, the cost analyses found that the average costs of SNF facilities were significantly higher than the average costs of ICFs. Of course, a rate setting method that uses level of care to group facilities and assigns higher rates to SNFs would be expected to perpetuate and magnify any differences between the groups. High cost ICFs would be unlikely to survive while low cost SNFs would be able to satisfy the SNF ceiling and could increase costs. Thus the significant difference in cost may be seen in part as an artifact of rate setting itself. It is reasonable, of course, to expect that in a world without public program payment requirements that SNF rates still would be higher than ICF rates since the higher classification of care requires more inputs which in turn involve higher costs. In this sense, it appears

---

[1] It appears that with regard to ceiling procedures that payment agencies, while willing to let facilities bear the losses entailed in being over the ceiling, are not willing to let them reap the gains of being under the "reasonable cost" screen.

that the cost difference uncovered by the cost functions estimation is a real one and indicates differences in the cost of two types of care that would appear in the absence of government intervention. Yet some state reimbursement practices with regard to SNF versus ICF rates exacerbate such cost differences. The typical procedure for calculating SNF and ICF rates is to consider the costs of homes in each classification level separately. While their costs may differ because of different input mixes, such stratification implies that their entire cost structures differ. However, the cost analyses found that the basic underlying cost structure of SNF and ICF homes was similar. An F-test rejected the notion that separate cost equations for each level of care were necessary; rather, a single cost equation was statistically appropriate which allowed only an intercept--constant--difference between the costs of SNF and ICF homes. This result suggests that states should investigate closely the validity of conducting separate cost analyses of SNF and ICF homes. By not grouping inefficient SNF homes with efficient ICF homes and vice versa, such separation may perpetuate inefficiencies in one or the other groups as well as adding extra administrative expense.

The following paragraphs review other typical grouping criteria and suggest some of the inefficiencies and inequities implicit in existing grouping methods.

Scale/Occupancy

The usual view here is that there are economies and diseconomies of scale and utilization with respect to facility cost and that they should be taken into account in determining the rate ceiling. Thus, if large (or small) homes cost more on average, then they can be paid more. This logic implies that the more efficient sizes of facilities are held to the higher efficiency standard appropriate to their size and are not rewarded for having chosen a lower cost scale of production. Similarly, the inefficient size homes are not penalized for their inefficiency. Thus, the firms in the industry face little incentive to move toward a more efficient firm size, if such exists. On an industry-wide long-run basis, such a grouping policy works against cost containment objectives by supporting production at inefficient scale.

There is an additional problem with scale and occupancy ceilings. The findings from the national cost data and from analysis of the three

states imply on the contrary that number of beds is not a statistically significant determinant of average operating cost, except for the smallest facilities. These findings were consistent across all four cost analyses, and it can be expected that where beds are used elsewhere as a grouping variable to set ceilings, differences between the ceilings for different size categories would be insignificant. We conclude that using bed size as a grouping criterion increases administrative cost in that it increases the paper work of the rate setting agencies, and works against unit cost containment by potentially supporting production in facilities of inefficient scale.

Similar arguments apply to occupancy ceilings. Here too, there was no evidence of occupancy rates having a statistically significant effect on operating costs. Thus, extra administrative cost is incurred at no gain in cost savings.

Perhaps the best that can be said about scale and occupancy ceilings is that they may help meet equity concerns and support maintenance of overall continuity of care for patients. Scale ceilings support the status quo size distribution of facilities and thus do not bring about the disruption that would occur when inefficient firms are driven out of business. Occupancy penalties serve to keep the stock of nursing home beds more full than they otherwise might be and so help increase access. However, if these equity concerns are the reasons for such ceilings, then it would be more appropriate to state them explicitly rather than overlook more direct and appropriate steps. For example, the little evidence found in the multivariate analyses for size effects suggested that very small facilities were more expensive per patient day than large facilities. Yet, when size is a grouping criteria, rate ceilings generally are higher for larger facilities. The logic appears to be that larger facilities tend to be newer and so have higher capital costs. But this logic would imply that size is relevant for capital cost reimbursement, not operating cost reimbursement.

Services

In general, there are two drawbacks to paying or developing ceilings on the basis of services. First, paying for inputs, by definition, creates an incentive to over-provide the allowable services beyond what

patients might choose themselves or need on the basis of medical/social condition.[1] This overemphasis on inputs is especially troublesome in that the inputs are not directly transferable to measures of output or performance. While it is easy to document inputs and their cost impact--the statistical cost analyses found that the service variables were important ingredients of variations in operating cost--little or no information is available on outputs. Since no state uses performance/output indicators in their rate determinations, homes have no reimbursement-related incentive to produce appropriate quality care.

The other drawback to paying for specific inputs or using input ceilings is that it has the potential to interfere with provider input decisions in a way that may be seen, on the one hand, as insuring high quality inputs and on the other, as preventing input substitution in response to relative prices. It is worthwhile to expand on this point. The input standards method is most often used for setting ceilings on nursing services. This is an input category where substitution between nurse aides, licensed practical nurses, and registered nurses is possible. If providers are given an overall dollar ceiling on the amount per patient day they can spend on nursing services, they can, if they wish, choose less costly input combinations; if the ceilings are set for each type of nursing input, cost-saving substitution is less possible. Management efficiency-type decisions are constrained. The number and specificity of the cost components that are limited by ceilings may be seen as the key determinants of the degree of

---

[1] That medically oriented services are typically more cost "allowable" than social or rehabilitative services is reflective of the predominance of the medical model form of care in nursing homes. While the study has no evidence one way or the other that this is due to reimbursement policy, in part that is because the data available were oriented around medical, not social services. Yet in itself this lack of social service data--especially since the state data were Medicaid cost reports actually used to calculate rates--is in itself support for the argument. Without cost data, social services can not be considered directly in the cost formula.

constraint placed on provider decisions about input substitution.[1]

Case Mix

It is assumed in case mix oriented studies of nursing home costs that more disabled patients receive more services and that therefore the cost of care to these patients should be higher. While no doubt there are costs to the additional services, more difficult patients do not necessarily receive more costly services in nursing homes even though they may require such services on the basis of need-based criteria. Rate setting methods which imply a willingness to pay more for more disabled patients must determine that they receive the services desired.

The major difficulty with this argument--and one which in part contributes to the absence of case mix concerns from reimbursement plans--is the difficulty of associating case mix and cost differentials. The problem here is controlling for the effect of quality. "Sicker" patients imply higher costs only if the care they receive in fact is appropriate to their "sicker" condition. Yet, the widespread absence of quality measures (in terms of outputs) means that statistical relationships between case mix or patient conditionand cost may be inaccurate. Indeed, this result occured in the statistical cost analyses where patient condition and costs were found to have little association. Consider the New York case. Initial cost analyses with no quality measures found no evidence of an association of case mix with cost. However, subsequent analyses which used a limited measure of quality

---

[1]Massachusetts at one time had separate ceilings for licensed nursing costs and other operating cost; any underrun on the operating cost ceiling could be applied to raise the nursing ceiling; but not vice versa, so that hiring more licensed nurses was encouraged. This was felt by the industry to be more constraining than a ceiling on all nursing salary expenditures, which allowed them to substitute among nursing categories more freely. In contrast, the non-profit providers, who were more costly and may be assumed to prefer a more costly nursing mix, preferred the licensed nursing ceiling method.

(from inspection reports) did reveal that non-ambulatory patients were more costly than ambulatory patients.[1]

In the national analyses, similar results were found: there was no consistent relationship between patient condition and costs. A compounding problem here on top of the absence of quality data was that the patient data were based on a sample of patients, so as to save costs in data collection. The result was that the cost estimates for patient condition were biased towards zero.[2] States that do decide to collect patient data and reimburse on the basis of them should be careful not to fall into the same trap, for while saving data collection money, the data will yield unsatisfactory results. One could conclude that it is not worth the effect to collect data on a small sample of residents, but that surveying all residents would be useful.

Input Prices

Input prices are used as discriminating variables so as to provide an equitable distribution of homes such that access is maintained across geographic areas (of both high and low input prices). This equity concern overrides efficiency concerns which would ignore input price variations that homes in a high cost area would simply go out of business, while those in a low input cost area would expand. This trend, of course, would be deterred to the extent that firms in a high cost area could reorganize themselves such that their productivity was higher. In fact, the state and national cost

---

[1] In the most detailed investigation of this issue to date, Winn and McCaffree (1976) also found no consistent association between patient condition, costs and absolute measures of services provided. One possible explanation for their results is that while they were looking at "best practice" homes, the homes come from a variety of states. Since standards of care vary widely across the nation, quality of care, even in best practice homes, also can be expected to vary.

[2] The reason is that errors in the measurement of an independent variable (when that variable is based on a sample, but the other variables are not) implies that the resulting coefficient parameters are biased toward zero (Johnson, 1972), i.e., there is a bias in the direction that patient characteristics have no effect on cost. The conclusion that the patient mix variables have little effect should not be made without some notion of the size of the bias.

analyses did find that input price measures were statistically associated with variations in operating costs. Thus, it is difficult to argue that any system of classification that takes into account such differences is particularly aimed at gaining efficiency. Rather, equity considerations widespread of availability of nursing homes creep into the rate setting process and are accepted by both the government and nursing home industry.

Given that rate setters do wish to consider input prices, they still must be careful in their measures of input price variation. One measure in effect over-adjusts for input price variation; when the facility's own average wage rate is used to determine its ceiling, higher ceilings apply both to facilities operating in high cost market areas (thus adjusting for price variation), but also to facilities choosing for some reason to employ a high cost labor mix or pay higher than the market wage. For example, in the national cost equation, it was found that a measure of wages for related, proxy occupations did affect operating costs. And when a measure of the wages actually paid to nurses in the home was entered, it too had a strong association with operating costs--although the coefficients on the remaining causal variables remained unchanged. If input price variation is to be allowed, adjustment for input prices should be made on a market area basis so that these cost-related variables remain exogenous to the facility's decision making process instead of coming under the facility's control. If not, cost pass throughs are allowed and, in fact, encouraged. Thus, it is important to choose an appropriate measure of input price variation exogenous to the facility. This implies that even measures of the average nursing wage in the county are inappropriate, for the aggregate combination of all the homes in the county exert control over the local average nursing wage. Instead, a broader measure of input prices should be used, say the average wage of service industry workers.

An alternative which is sometimes used to allow for input price variation is the use of region. The problem with geographic measures of price variation is that if rural facilities were less costly because of better management or more efficient building design, they would in effect be penalized, because these savings would be treated as if they were due to lower factor prices. Any method is open to this criticism if it introduces

region into the rate setting formula rather than using factor price differences explicitly.[1]

Ownership.

For-profit/non-profit status is classically used in rate setting discussions to dichotomize homes in terms of services of level and quality. While all the cost analyses conducted for this study found that non-profit homes had higher costs holding many other factors constant, these other factors included a broad range of services. Thus, one can conclude that the costs of non-profit homes are higher for unmeasured reasons--either higher quality or inefficiency. Conclusive tests of these possibilities were not possible because of the absence of quality data. However, non-profit facilities in New York still had higher operating costs even after controlling, in a rough way, for quality. While conclusive statements about their inefficiency can not be made, no evidence available to this study supported the proposition that non-profit and for-profit homes had cost difference for reasons of product differentiation. It is our best judgement that rate setters that differentiate on the basis of ownership are creating incentives for inefficiency.

---

[1] A multivariate determination of the impact of region may be preferable, since other cost-related factors are presumably held constant leaving input prices to be explained by the impact of the regional variables. However, when regional prices change at different rates within a region, it would be more effective to account for the impact of price explicitly rather than dealing with the net impact through a regional grouping mechanism.

TECHNICAL PAPER #3

Technical Paper #3

METHODOLOGICAL AND ECONOMETRIC APPROACH TO COST ANALYSES*

T-3.1 Overview

The purpose of this technical paper is to set forth the methodology for the development, estimation, and implementation of nursing home average operating cost models on both the national and state-specific level. The actual tactics of the general strategy laid out here are couched in the context of the cost analyses reported in Technical Papers T-4 to T-7 of this study. Since the discussion there as well as in Chapter 3 reviews many of the issues involved in cost function estimation, particularly with regard to variable specification and hypothesis testing, such discussions are not repeated here. Rather, the discussion here is with more methodological concerns that might be but should not be overlooked by rate setters interested in performing cost analyses on data from nursing homes in their state.

Section T-3.2 identifies data needs for both national and state-specific cost modeling. The discussion there draws heavily from both a conceptual model of nursing home cost behavior developed in Technical Paper #1 and our actual experience in the modeling of nursing home costs using four empirical data bases: a National NCHS probability sample, and Medicaid cost report data from the States of New York, Massachusetts, and Indiana. Section T-3.3 discusses the tactics of data "cleaning," i.e., the steps required to weed out bad data values. Here again, the discussion draws on the specific measures taken in our cost studies to insure the reliability of the resulting cost analyses. Sections T-3.4 and T-3.5 identify econometric issues of model specification and estimation, respectively. These sections are quite general in that the issues discussed are not unique to researchers of nursing home costs. Section T-3.6 concludes the paper with a discussion of the rate-setting role and the strengths and weaknesses of the regression analysis approach to cost function estimation.

T-3.2 Data Needs

T-3.2.1 Specifying the Nursing Home Product

A cost model may be of great use in designing or maintaining a reimbursement scheme which first attempts to screen out facilities which are

---

*By Gail Jensen, Howard Birnbaum, Christine Bishop, and A. James Lee.

excessively high cost, and second, pay the cost of average and efficient production. The first step in the development of empirically based cost functions is the specification of the product which rate setters are willing to pay for. What is it that we want to know:

- The cost of using specified resources (such as nursing hours, administrative hours, supplies, and medications) which are used in the process of providing long-term care? or
- The cost of producing certain outputs or outcomes which are characteristic of the process of providing care (such as quality of care, services, and patient-days provided)?

It is answers to the second question that are crucial in that it determines what we will measure the costs of. In the cost analyses of this study, the basic unit of output is taken to be the patient-day. However, the cost of a patient-day will vary by services provided. As these services are important distinguishing factors in cost determination, they must also be included in a cost model.

The number of admissions per year, patient debility mix, facility age, certification and scale of plant are also relevant in cost determination and will be included in a model which explains cost.

A general model of total nursing home operating costs might look as follows:

$$C = g(Q, A, S, X, P, R, Y) \qquad (1)$$

where: C = annual total operating costs

Q = total patient days provided

A = number of yearly admissions

S = index of types and quantities of services provided (e.g., number of patient-days for which oxygen therapy was provided)

X = patient debility mix indicator

P = vector of input prices

R = state regulations

Y = vector of other facility characteristics such as profit non-profit, age of facility, scale of plant (function of patients, rooms, beds), rural-urban location, etc.

Discussion with regulators at the state level has indicated that equation (1) above is one of the most useful forms of the cost function from a policy standpoint. That is, a function which disaggregates nursing home outputs (the patient-day) into services provided permits comparison of cost at various service levels--a useful input into policy decisions regarding the quality which public payors are willing to provide. However, entrepreneurs would still be free to combine inputs as they wished in providing these reimbursable outputs.

Policymakers would need to make several decisions even given an estimated version of equation (1). Among them are: which services are reimbursable (e.g., choice of meals) and how can across-patient differences in service requirements be estimated. That is, provision of the same reimbursement across all patients would provide incentives to favor less resource-intensive applicants, an undesirable outcome. Some screening procedure must be devised.

Thus, the cost function may provide the link between costs and output (patient-days and service levels per day), but not between patient characteristics and required service levels. That is, quality of care provided (given services) to a person with particular attributes is necessarily a policy-choice variable.

The pure cost function of economic theory assumes rational economic behavior. Its building blocks are similar production possibilities across firms, cost minimizing behavior, known input prices and target outputs. In practice these assumptions may not always hold the pure model must be adapted to actual characteristics of the nursing home industry. We now turn, briefly, to these amendments.

T-3.2.2 The Specification of Variables

This section and the remainder of the paper maintain that outcome measures are quantities that one wants to examine. In the following paragraphs we identify the types of variables that should be considered in measuring the output of nursing homes.

T-3.2.2.1 Output Variables

Skinner and Yett (1970) have shown for nursing homes that case flow measures of output affect costs. Because increased case flow involves

additional administrative work, homes with greater patient turnover, shorter lengths of stay, and greater admissions rates are likely to have greater costs. Case flow variables also can be viewed as describing patient care provided; a short-stay patient may be more likely to require intensive and rehabilitative care than a long-stay patient using more maintenance services.

However, unlike many other industries, output in the nursing home context is not a unidimensional quantity. While the basic unit of analysis for nursing homes is typically taken to be the patient day, it is more appropriate to consider multiple measures of output. The number of patient-days provided must be augmented to reflect the "quality" of output as well as patient characteristics, specifically the health status and debility of a patient. For example, a mix of patients primarily unable to feed themselves may require more nursing hours per patient day, with commensurately greater costs, than a mix of relatively self-sufficient patients. Hence, indicators of patient status (i.e., case mix) should be included in the cost model. Similarly, costs may vary by the quality of the output produced. The difficulties in measuring quality in health care research are well known. To the extent possible, quality is taken into account by incorporating service variables in the model, such as the presence or absence of occupational therapy, the percentage of one-bed rooms, etc.[1] Very little quality-specific information was available in the data used here (or in other nursing home data); its absence and the use of services as proxy variables raise issues of interpretation since one cannot distinguish between services which are inappropriate to patients or inefficiencies. In any event, one output of nursing home is a vector of characteristics including patient-days, as well as case mix, quality, and services.

T-3.2.2.2 Measuring Patient-Mix Attributes

Case mix will appear in the nursing home's cost function and in the equations which measure the public and private composition of the patient-days supplied by nursing homes, so it is important to be able to define and quantify a measure of case mix. For the purposes of this study

---

[1] Technical Paper #1 discusses how the union of information on patient mix and services provide a rough approximation of "quality."

the index used should first, relate to the overall patient mix of debilities in each home, second, capture the cost related aspects of the debility mix, and third, eliminate the problems found in the earlier attempts of researchers to aggregate the spectrum of a patient's debilities into a useful index. With respect to relating to overall patient mix, the index should be designed so that it reflects both the range of disabilities the nursing home must deal with and the percentages or intensitives of various disabilities present in the home.

The problem of constructing such an index for each nursing home is different from but related to the problem of measuring a debility index for a particular patient. There is a body of literature dealing with the latter problem and a range of patient debility indicators have been developed. For several reasons it is not clear that the literature is useful. Most important, in the bulk of existing literature, little attention is paid to the cost related aspects of debility. The three studies that do (Skinner and Yett, 1973; Muller et al., 1963; and Watson and Fulton, 1967) find a weak relationship between the index developed and nursing home costs. We feel that this finding is the result of poorly constructed indices.

Overall, the types of indices developed have the problems of first, not utilizing all information pertaining to the range and the composition of patient debilities; second, not acknowledging the fact that different levels of various disabilities have different amounts of nursing care (hence costs) associated with them; and third, relating to the debility level of particular patients rather than to a group of patients.

The development of a debility index for the home, as for the patient, must focus on several issues. First, what is debility, i.e., what should it measure? Should it measure the need for timely assistance from nursing home personnel in order to be able to perform normal activities of daily living (ADLs)? Or instead, should it measure the degree to which an individual's mental and bodily functions are impaired irrespective of which bodily functions have a significant impact on the patient's ability to perform ADLs? The former is directly related to the cost of maintaining a patient since more assistance from nursing home personnel is more costly. The latter is not necessarily related to costs. For example, an individual who has sustained a Cerebro Vascular Accident (CVA) might be considered normal and independent

in the performance of various tasks of day-to-day living in a nursing home although his body can be considered impaired if portions of brain tissue have been completely destroyed. The residual effect of his CVA, namely, non-functional portions of brain tissue, although an impairment in a medical context, should not be defined as a disability in constructing a useful cost related debility index.

After deciding what it is we want to measure, the second question becomes: how will it be measured and is a one-dimensional index adequate? To answer this question, we initially look at answers provided by the literature. The indices appearing in the literature attempt to answer the questions above in a manner which best accommodates the analyses in which the indicators are used. For purposes of this study all are deficient.

To a large degree, the debility associated with a particular patient's case can be represented by the extent to which various activities of daily living (ADLs) are constrained. The level of ambulation, bowel and bladder continence and control, speech, vision, hearing, mental status, feeding ability, muscular strength and coordination, ability to dress one's self, and bathing ability are typically accounted for in some way when constructing the patient's debility index. There are differences in the literature with respect to which ADLs are used in the construction of an index, the method used to represent information on a patient's performance in each ADL, and the technique used to aggregate the data on a patient's performance in each ADL.

Measures of Debility Found in the Literature

The first general type of index to be examined can be referred to as a sum-of-scores index. The index is ordinal, so that for each ADL (feeding, bathing, communication, dressing, use of household devices, etc.) a patient is assigned a score reflecting his relative ability to function independently in that activity. However, because the score is ordinal, the range of the interval between successive scores are not of equal value in terms of some absolute concept of debility. For each ADL, the criteria used to establish which score corresponds to each level of performance is determined prior to patient evaluation. In this way, different examiners are supposed to assign similar scores for the same patient. Complete

independence in a particular ADL corresponds to a score of zero for that ADL and complete dependence corresponds to the maximum score on the scale. Degrees of partial disability in an ADL are measured by intermediate scores on the scale. The grid of the scale used to assign scores for each ADL varies across the authors who have used this type of index. The arithmetic sum of a patient's scores across the ADLs being examined is taken as the index measure of his/her overall debility.

Authors employing an index of this type include Carroll [1962], Wylie [1964], Dinnerstein, et al. [1965], Linn [1967], and Schoening and Iversen [1968]. The index has been used to evaluate patients and patient recovery potential.

Unfortunately, the sum-of-scores index does not account for differences in the amount of care (hence costs) associated with different performance levels of the various ADLs. Thus, implicitly, ADLs are treated equal in the degree of care associated with them. Knowledge of an individual's sum-of-scores does not provide information about which debilities generated the score. For our purposes, this is clearly a deficiency since both the range of disabilities and the types of disabilities present are related to the costs of caring for the patient. Moreover, it is possible that the interpretation of the criteria used for patient rating may not be consistent among examiners within the same home or across different homes. Inconsistent criteria interpretation could result in the wrong classification of some patients. For this reason the index may not be reliable.

A special case of this general type of index can be referred to as the Yes/No Sum-of-Scores Index. In constructing this index, for each ADL a patient is considered either dependent or independent in the activity. The cut-off point for independence is usually selected in an ad-hoc fashion. Independence corresponds to a score of zero and dependence to a score of one. There are no intermediate values. The total possible sum-of-scores equals the number of ADLs in which the patient is dependent, so score interpretation is straightforward. Having a score of five corresponds to possessing any five disabilities.

The Yes/No index has several problems. The index is not designed to be cost related, nor does knowledge of a person's score tell us for which ADLs the person demonstrates dependency. As a result different event scores are associated with different costs of care. For example, a score of 2 could indicate either dependency in feeding and continence or dependency in communication and transferring from a bed. The first pair of disabilities may correspond to considerably higher costs of care than the second pair of disabilities. Also, partial dependencies in ADLs have associated costs, but are not accounted for in this index.

An example of the use of this index to evaluate patients for the purpose of determining staffing requirements can be found in Lawton and Brody [1969].

Another debility indicator, called a Guttman index, has been used by Skinner and Yett [1973].[1] They examined the relationship of the index to costs, care needed, and duration of stay. The Guttman debility index is identical to the sum-of-scores index in actual construction but different in terms of the way scores are interpreted. Use of the Guttman debility index is based on the assumption that the probability is very high that as a patient changes from being completely independent to completely dependent, he will do so by becoming dependent in the various ADLs in one pattern only. Put another way, suppose that a total of n ADLs are being examined. Then this procedure implies that out of the n factorial possible patterns for successively incurring disabilities, the probability is very high that investigators will find one and only one pattern common to all patients. A deducted assumption is that patients with the same scores will have a very high probability of possessing the same types of disabilities. For example, in a set of 5 ADLs, we would expect that all individuals who exhibit depen-

---

[1] Skinner and Yett used a simple dichotomous categorization scheme (either independent or dependent) to classify patient performance on each ADL. There is another type of Guttman scale different from that used by Skinner and Yett allowing for more than two 2 categories of classification on each ADL. (For a discussion of this index see Edwards, A.L., Techniques of Attitude Scale Construction. New Jersey: Prentice-Hall, Inc., 1959). The second type of Guttman index has not been used to measure debility but has many of the problems associated with the Skinner and Yett index. It will not be used on the cost studies.

dency in 4 ADLs do so with respect to the same 4; that those who exhibit dependency in 3 ADLs do so with respect to the same 3, and that furthermore these 3 are among the 4 dependencies exhibited by those who are dependent in 4 ADLs; that those who exhibit dependency in 2 ADLs do so with respect to the same 2, and furthermore that these 2 are among the 3 dependencies exhibited by those who are dependent in 3 ADLs, and so on.

The problems with the Guttman index are first, that it is unrealistic to presume that we can know which debilities have most likely generated particular scores. Second, the amount of care associated with different disabilities has not been accounted for. Third, and most important, it is unreasonable to presume that patients with the same number of disabilities will have the same type of disabilities.

A type of multidimensional index has been employed by several authors. The index score for each patient is represented by a vector of length equal to the number of ADLs being examined. Each coordinate of the vector is a score reflecting the patient's degree of independence in performing one of the ADLs. A score of zero in coordinate i reflects complete independence while the maximum score, say four, reflects complete dependence in performing task i.

The criteria for evaluating a patient's degree of dependence and assigning a score to his dependency level in each ADL is common to all patient evaluators. Although this index is ordinal, it provides information on both the range and type of disabilities demonstrated by a patient.

Examples of analyzing patient debility in this way can be found in Muller et al, [1963], Watson and Fulton [1967], Scranton, et al. [1970], Schoening and Iversen [1968], and Gersten et al. [1970]. The index has been used to estimate staffing expenditures and the typical time pattern of patient progress. It has also been used to determine the relation of scores to strength and range-of-motion indicators. By presenting a matrix of debility scores, this index does resolve the problem of aggregating various disabilities into one index.

Another type of index, employed by Young [1962], Katz [1963], [1970], and Gordon and Kohn [1966] can be referred to as a categorical

index. This index permits a patient to be classified as belonging to one of several possible debility categories. Each category corresponds to specific patient evaluations on selected ADLs. The patient's overall debility category is typically determined by looking at his performance in various tasks simultaneously. Debility categories might be labeled A, B, C, D, E, F, and "other." Category A reflects the most independent type of patient; categories B, C, D, E, and F reflect successively more dependent patient types. This type of index has been used to determine patient recovery potential and nurse staffing requirements.

This index, although ordinal, succeeds in providing information on both the range of disabilities and types of disabilities present in a patient. In this sense, it accounts for the cost related aspects of different debility levels. However, there is no accurate method to determine the costs associated with the "other" category. This presents a problem when a large proportion of patients examined fall into this category; as in the Katz study where 14% of the population fall into the "other" category.

The indices discussed briefly above, even without their associated problems, are not directly applicable for the purposes of this study since they deal with the evaluation of individual patients rather than homes. The problem of deciding on an appropriate index which reflects a nursing home's debility mix is examined below. Three alterntive indicators of aggregate patient debility are discussed with attention paid to their suitability for use in a cost-function.

Alternative Debility Measures

Measure 1. The first nursing home case mix measure is a natural extension of the one dimensional indices used to rank patients in terms of debility, discussed previously. However, the index becomes two-dimensional, by providing both a measure of the average patient level of debility and the

range of dispersion of patient debilities found in each home.[1] Use of the index presupposes that patients in all homes have been evaluated using the same index and the same set of debility level standards. It is assumed that the index for patient debility evaluation is one of the one-dimensional indices presented above.

Although such an index has intuitive appeal it provides only ordinal measures of debility, i.e., the range of the intervals between successive scores are not of equal value in terms of debility. Therefore, we cannot measure the additional care required (hence additional costs) associated with changes in debility. Interpreting the values of estimated coefficients after cost function parameter estimation is unclear.[2]

---

[1] The index measure for home k can be represented as:

$$X_k = (\hat{DI}_k, \hat{Var}\ DI_k)$$

where

$\hat{DI}_k$ = measure of the scores' central tendency for home k such as the mean, median or mode of patient scores.

$\hat{Var}\ DI_k$ = measure of the dispersion or variation in patient disabilities for home such as

$$\frac{1}{M_k}\sum_{i=1}^{M_k}(DI_i-\overline{DI})^2 \text{ or } \frac{1}{M_k}\sum_{i=1}^{M_k}|DI_i-DI_{Median}|$$

where

DI_i = debility score for patient i in home k

M_k = number of patients in home k

$\overline{DI} = \sum_{i=1}^{M_k} DI_i/M$ = mean score of patients

DI_{Median} = median score of patients

[2] Investigators who have used a one-dimensional patient debility index in cost analysis have found that patients with mid-range scores are more costly than patients with either very high or low debility scores. This suggests in reference to an earlier footnote that DI should enter the cost function in a quadratic form,

$$\beta,\hat{DI}^2 + \beta_2\,\hat{DI} + \beta_3 \quad (\beta_1 < 0,\ \beta_2,\ \beta_3 > 0).$$

There are other problems with Measure 1. In a home's score (as in a patient's score) there is no way to detect which types of disabilities are present.[1] It is important to account for disability types since types will vary in the required nursing care (hence costs) associated with them. For example, a hearing disability will probably not require as much attention as a locomotion disability, although a home with a score of 1 could reflect an average of either disability type across patients. Two homes with equal index values may have quite different costs depending on which patient disabilities generate their score. Also, it may be naive to suppose that the basis for X patient classification is identical across homes for which the index is used. Debility criteria may be consistent within nursing homes but inconsistent across homes. In this case, comparisons between nursing homes become less meaningful.

Measure 2. Another approach is as follows. For each home the percentage of patients possessing each level of each debility is recorded. Then an estimate of patient day nursing hours required to care for each level of each debility is obtained. The estimate of nursing hours required could be constructed by averaging the opinions of a group of consultant nurses or by choosing the median opinion.

The debility indicator for home k is taken as

$$X_k = \sum_i \sum_j f_{ijk} H_{ij}$$

where

f_{ijk} = percentage of patients in home k with level i of debility j

H_{ij} = estimate of patient day nursing hours required to care for level i of debility j

Upon inspection one sees that

X_k = (total number of nursing hours in home k which should be allocated to all debility needs)/(number of patients in home k)

≈ number of nursing hours which should be allocated to debility needs for the average patient in home k.

---

[1] The Guttman scale is an exception.

This index has the advantage of being cardinal and capturing the cost related aspects of disability types and levels of each type.

Three potential deficiencies are appartent: first, the dispersion of disabilities is not accounted for and the size of the dispersion of debility types may be positively and significantly related to operating costs; second, it is likely that many investigators will question the validity of estimation results because of disagreement on our choice of values for the H_{ij}'s, nursing hour weights, regardless of the method used to obtain them; third, the criteria used for patient debility level classification may be inconsistent across homes.

Also, as with the two-dimensional index, knowledge of a home's score does not provide information on which debilities generated the score, but since the cost related aspects of disabilities taken into account, this information is less crucial.

Measure 3. A third approach is to use obtained data on the percentages of patients in home k with different levels of the various debilities as a measure of the debility mix for home k. Letting f_{ijk} equal the percentage of patients in home k with level i of debility j, the f_{ijk}'s could be entered directly into the operating cost function.

This third approach has the following advantages. The f_{ijk}'s are ordinal rather than cardinal. Information on both the types of debilities present in a home and the range of disabilities present is contained in the measure. Letting f_{ij} equal a column vector with coordinates f_{ijk} (k = 1,2, ...n where n is the total number of homes used for cost function estimation), interpretation of the estimated coefficient of each f_{ij} is relatively straightforward. The estimated coefficient of f_{ij} is an estimate of the average increment to operating costs resulting from a unit increase (a unit being a percentage point) in the proportion of total patients possessing dependency level is of debility j. The costs associated with various combinations of debilities and debility levels are measured directly. This contrasts sharply with the two approaches presented earlier which do not provide any way of allocating costs.

A potential problem exists in using this third debility mix indicator, namely, the criteria used for patient debility level classification

may be inconsistent across homes. However, the possibility of inconsistent interpretation of debility level criteria is present in every index which is available and discussed here. Apart from this potential deficiency, use of the third nursing home debility mix index appears advantageous relative to the first and second nursing home indices discussed. Hence, as a measure of cost-related debility in this study, we entered the f_{ijk}'s directly into the cost function. The f_{ijk}'s are cardinal. This property facilitates a clear interpretation of estimation results. Note that in using the ordinal two-dimensional index there was no meaningful interpretation of estimated cost function parameters. With the choice of the third index, the estimated coefficient of each f_{ij} represents an estimate of the average incremental costs resulting from a percentage point increase in the proportion of total patient possessing dependency level i of debility j.

With this choice of index, the combination of debility levels and debility types was accounted for. Neither the first or second alternative measure provides this information. As emphasized throughout this section, it is the type of debilities present in a patient or home that influence costs. The third index also accounts for the range of disabilities each home must contend with. This property, lacking in the second nursing home index examined, enables us to capture the costs associated with the degree of dispersion of debility types present in a home.

In addition to this ADL index, our cost studies explored the usefulness of patient age, sex, and diagnosis descriptors. (The reader is referred to Technical Papers #4 through #6 for detailed descriptions of what was actually done with the available data bases.)

T-3.2.2.3 Measuring Characteristics of the Regulatory System

Because the nursing home industry is regulated, a second set of modifications to the traditional cost functions of economic theory are required. In an unregulated industry, facility type need not necesarily be distinguished to the extent that relevant production differences are captured by the patient and service characteristics. However, since states typically impose differential requirements (e.g., nurse staffing requirements) and may explicitly consider facility type in establishing reimbursement rates, differential regulatory requirements suggest extraordinary cost differences

for two reasons. First, as a consequence of regulation, historical cost differences may be maintained over time independent of patient status, "quality" of care, etc. Second, regulations may impose cost differences because of differential resource requirements, as for example, by imposing staffing requirements. Therefore, it is necessary to take regulation and reimbursement practices explicitly into account.

One of the most difficult tasks in the estimation of a nursing home cost function is to choose the appropriate way to include the effects of regulations. Regulations affect the firm's decision making process and therefore influence the cost function revealed by the data. Furthermore, we are interested in the manner that nursing home operators alter their behavior in response to regulations. Therefore, in a cross sectional analysis, state to state differences in regulations must be accounted for. The cost function that is relevant for the design of an optimal prospective reimbursement system is one in which the variation in costs due to regulation is parceled out and absorbed by variables representing regulatory effects. The omission of regulatory variables means that their effects cannot be determined and that other coefficients may be biased if these variables and the regulatory variables are correlated. Consequently, the appropriate inclusion of regulatory variables is a key issue in the specification of a cost function.

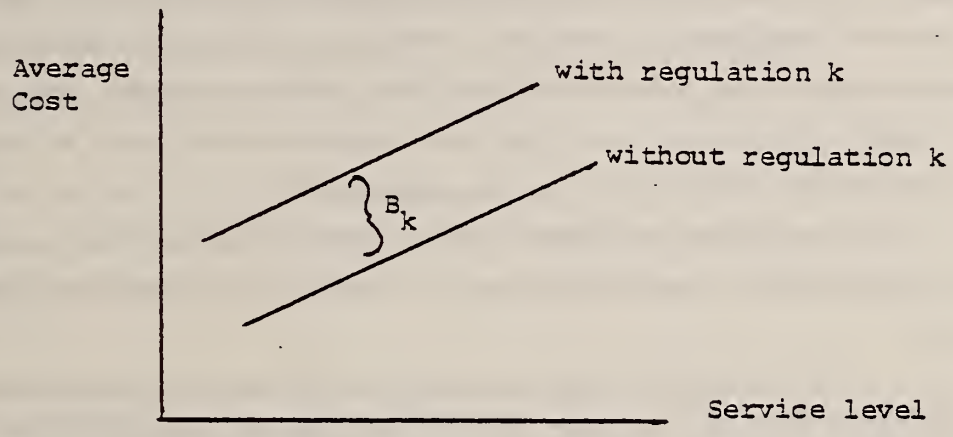
Available Techniques for Estimating Regulatory Effects

In recent years, several empirical studies have been made of the cost, production and investment functions in regulated industries. A review of some of the relevant literature will provide insight into the standard techniques for the econometric modeling of regulation. Sloss (1970) undertook an econometric study of the effects of regulation on the rate for moving goods per ton mile in the ten Canadian provinces. After a careful documentation of province to province differences in regulations, he roughly divided the provinces into two groups and used a dichotomous variable representing regulation versus no regulation to deal with cost variations due to structures. The approach adopted by Jorgenson and Handel (1971) in their analysis of investment behavior in U.S. regulated industries, marked by naturally increasing returns to scale, perhaps best illustrates the econometric techniques used

to model regulation: "The effects of the regulatory constraint are represented by the difference between the regulated price of output and its opportunity cost. If regulation is effective, the actual price and output are the same as would prevail under a competitive market structure. In this study we assume regulation is effective - that is, that regulatory price equals opportunity cost." (Ibid. p. 217). Similarly, Spann (1974) in his application to a production function for electricity of a test of the Averch-Johnson thesis that overcapitalization occurs in regulated industries where the rate of profit is tied to capital stock completely ignored state to state differences in regulation and did not include any regulatory variable. He was, however, aware of wide state to state variations and noted that two states, Texas and Minnesota, did not regulate at all on a state level. In a similar article dealing with electric utilities, Courville (1974) also failed to include a variable representing cross-sectional variations in regulation, although he remarked on the complex and widely variable regulatory structure.

The most ambitious effort to build variations in regulatory effects into an econometric study of a regulated industry was undertaken by Peterson (1975). He tested the Averch-Johnson thesis on a cross-section of fifty-six steam generating plants. He writes: "Tightness of regulation proved difficult to quantify." He uses three dummy variables: a dummy variable that takes on a value of one in states with a regulatory commission and zero in states without commission; a variable that takes on a value of one in states which base their rates on an original cost basis and a value of zero in states that set rates on the basis of fair value or have no commission; lastly, he developed a continuous variable ranging between a value of zero and one which captures the effect of a liberal or conservative commission on returns to equity capital. The fairness dummy variable was statistically significant at the 95 percent level, while the other regulatory dichotomous variable was not. This can perhaps be explained by collinearity between the variables. The continuous measure of the effects of a commission in returns to equity capital was not only statistically significant, but also it was employed in a test of the Averch-Johnson thesis; that test confirmed the theorem.

Several econometric techniques are available to control our national model for the wide across-state variation in the regulations that govern nursing homes. The first is similar to the method employed by Peterson (1975). Here, the presence or absence of a particular form of regulation (e.g., minimum capacity regulation, required service offerings) might be designated by a dummy variable in a regression equation. The estimated coefficient would reflect the average addition to cost from some regulation, K, as indicated below.



Data relevant to a cost model would include both legislative and enforcement variables. The former include the type of Medicaid reimbursement scheme (broadly defined) minimum input requirements (nursing hours per patient day) and certificate of need constraints. However, resorting to written regulations would only be begging the issue of regulatory effects. Because of differences in enforcement, the simple inclusion of variables representing written rules in the cost function would not produce a cost function devoid of regulatory distortion. Unfortunately such enforcement effects are hard to quantify under any circumstances. The absence of data on these enforcement effects means that they cannot be taken into account in the cost functions. Hence we must be careful in interpreting our results and present appropriate caveats.

Turning to the issue of legislative/regulatory variables again forces one to confront data problems. Regulatory data specific to individual states has not been centrally codified or even collected under reliable

conditions for 1972, the year of the principal data base for the empirical work, the NCHS national housing home survey. Fortunately, there do exist several data sources which are relevant to the problem, although each has limitations. In our study we attempted to piece together a set of variables which proxy some (but, unfortunately, not all) of the regulatory information which is desirable.

Estimation and Control of Regulatory Effects

Regulatory effects had to be taken into account for two reasons. First, it is desirable to have quantifiable estimates of the effect of different regulatory variables. Second, because the regulations differ across states, to the extent that they influence costs, the regulations must be taken into account when pooling cross-sectional data so as to avoid biases on the other coefficients in the equation.

Including accurate data on regulatory variables solves both concerns. Our first effort, therefore, was to work with the available regulatory data.

The available data sources for estimating (and controlling for) regulatory effects included Applied Management Sciences (AMS), General Accounting Office (GAO), Urban Institute (UI), and Battelle. They faced a common set of problems. First, they were not designed for estimating the effect of reimbursement practices on nursing home costs. Many of the variables of interest to this study were not included: e.g., what grouping criteria were used? What cost variables had ceilings and what were the levels of the ceilings? What was the required occupancy level? What input standards were in effect? etc. Second, the answers to the questions that were asked relied on state interpretation of the questions rather than their choice among a rigorous formulation of possible answers. This problem was especially troublesome in evaluating the payment mechanism answer. That a state said it had a prospective reimbursement system still leaves much ambiguity since "PR" can mean a multitude of things. Third, with the exception of the Urban Institute data and perhaps some of the Battelle data, the data were for years other than 1972. The first two problems simply reduce the richness of the

analysis which can be undertaken. The problem of timing can be resolved.[1]

Payment mechanism results for the year in question (1972) are bracketed by the AMS data for 1973 and the GAO data for 1970. The UI had data for 1972. We used the AMS and GAO data as verification of the UI data. While the sources answer slightly different questions, it was possible to merge the answers into a unified format.[2]

In addition to using the above "merged" data in an explanatory variable in the cost functions, one could estimate, as well, separate cost equations using the "raw" data of the AMS and also the UI sources. Although this was not done in our study, it would be an interesting analysis.

Data on other regulatory variables were also available from the secondary sources.

If it had turned out after preliminary results were in that the regulatory data were not useable, an alternative approach would have been to employ 50 state-specific dummy variables. This would act as a control for

---

[1] An initial attempt was made to resolve these problems by investigating the likely results of conducting our own telephone survey. After a few explanatory phone calls, it became clear that (a) it was difficult if not impossible to locate individuals knowledgeable about 1972, and (b) reliable answers to the questions we wished to ask were not possible. Memories and records do not appear to span the period between 1972 and 1977. As a result, the Project Officer for this study directed that we stop such efforts. Hence, this study will rely on existing data.

[2] The decision hierarchy for resolving differences among the sources was:

1. If all three sources agree that the rates were cost based, we set R = 0; if prospective, we set R = 1; if flat or point system, we set R = 2.
2. If the UI and AMS data agree, but the GAO data disagree, we accept the UI/AMS results, since they are for the adjacent years of 1972-73.
3. If the UI and AMS data disagree but the AMS and GAO data agree, we will accept the AMS/GAO results. (Since the 1973 AMS and 1970 GAO results bracket the 1972 UI results, we would argue that consistency over time dominates the discrepancy in the interviewing year.
4. If all three sources disagree, we make a phone call to the state in question.

An additional procedure was to review the AMS discussion of the state reimbursement procedures; some of their data categorizations did not reflect their verbal discussions, but corroborated UI results.

state-specific characteristics for which we did not have measures. The disadvantage of this approach was the large number of additional coefficients which must be estimated (49). Under these circumstances, problems of multicollinearity were likely to become severe and many degrees of freedom would be lost diluting the power of our statistical inferences. Moreover, the cost of inversion of large matrices further reduced the desirability of this rather crude approach.

T-3.2.2.4 Other Explanatory Variables: Input Prices and Facility Characteristics Describing Production Techniques

In addition to output and regulatory variables, the cost analysis must consider two other types of variables. The first includes measures of input prices. The last vector of variables which may affect nursing home characteristics includes an ad hoc set of facility characteristics. The facility characteristics reflect alternative methods for producing comparable output, and may reflect, for example, managerial decisions and abilities to operate on the general cost curve facing all nursing homes.

T-3.3 Data Cleaning

It is inevitable that some facilities will have incomplete or erroneous records. It is also likely that the researcher will find some facilities which he (she) believes to be non-representative for purposes of examining nursing home cost behavior. (For example, facilities which expanded or contracted their bed size by more than 10 percent may be of interest for a study of nursing home expansion, but non-representative for determination of ongoing average costs.) Below, in the context of the national analysis we identify alternatives to the simplistic approach of "throwing away" observations which exhibit (1) any missing observations on the variables of interest or (2) any unintuitive data values on the variables to be studied.

In checking for obvious errors, the researcher must examine the mean, range, and variance of each available variable and inspect for errors which can be detected in a straightforward manner such as a counter intuitive mean or extreme outliers (very high maximum or very low minimum). Also helpful in the detection of bad data values is the construction of histograms for each variable to be used for analysis.

Of key importance is the nature of missing observations. How are missing values coded? How many homes report the variable to be modeled, operating expenses? For homes reporting this variable, how many are missing data on the scale of production? The scale of production and services? These and patient mix descriptors? etc. . . If an observation is missing only a few of the data points, then one pragmatic solution is to approximate the missing data points by the mean values for the corresponding variables for similar types of facilities, e.g., if the home is a non-profit home in county X of bed size Y, and is missing its occupancy rate, then the mean occupancy rate for non-profit homes in county X of bed size Y can be substituted. However, if many variables are missing, the observation probably should be deleted. The exception to this procedure involved homes missing the dependent variable, average operating costs. Such homes always should be deleted. Also important here is the nature of multi-collinearity among variables of interest. By constructing a correlation matrix and inverse correlation matrix for the variables under study, one can measure each coefficient of determination, R^2, associated with the least squares regression of a particular predictor, X_j, regressed on the remaining (j-1) predictors.[1]

To illustrate data cleaning procedures, the following paragraphs discuss in detail what data cleaning activities were undertaken for the national cost analyses in Technical Paper #4.

For the 1569 homes reporting the dependent variable, average operating costs, the following decisions and observations on the nature of the data were made.

---

[1] One criterion, suggested in the literature on multi-collinearity, for the assessment of multi-collinearity is that the variance inflation factor of each variable be less than 4.0. (See Fisher (1977) for this and other detection tests.) The variance inflation factor of each variable, X_j, is given by

$$C_{jj} = 1/(1-R^2)$$

where C_{jj} = the diagonal element of the standardized $(X^1X)^1$ matrix corresponding to the jth variable, X_j, in the data matrix, X. This procedure is commonly called the Ferrer-Glauber method.

Ninety homes had a value of the occupancy rate greater than one. Some of these homes have been kept on the sample since there are at least two possible reasons why occupancy may have been greater than one. First, if a home is running close to capacity throughout the year and realizes a decrease in the number of beds shortly before completing the questionnaire, occupancy will likely be greater than one. Second, if the home was running at near or full capacity throughout the year and frequently took in more patients than it had regular beds for, occupancy could be greater than but close to one.

However, four homes had a value of occupancy greater than a specified cutoff, e.g., 1.5. These homes were eliminated on the somewhat arbitrary grounds that while occupancy rates greater than one are possible, the previous reasons were unlikely to justify such a large occupancy rate.

One home reported a bed count of 3252. This value was replaced by 325 since 325 is consistent with other data reported by the home (e.g., ICF beds of 14, SNF beds of 180, census of 3180). Another home which reported ADMIS and DISCH of 1751 was deleted from the sample since this extreme outlying data did not seem compatible with other reported data (e.g., census of 91, beds of 108, certified for Medicare and SNH, patient days of 32409).

Eighty-seven homes have beds which are certified as both SNF and ICF beds. For such homes, beds which are certified for both were considered as SNF since SNF certification has more stringent regulations associated with it.

Thirty facilities, each best described as an extended care unit of a hospital, were deleted from the sample used for cost function analysis since such facilities are essentially different from nursing homes in terms of the type of patients treated, the scope of services provided, and operating costs.

It was not clear whether or not homes which were not certified as SNF, ICF, or Medicare, or some combination of SNF, ICF, or Medicare should be included in the sample used to estimate a cost function which will be used in the reimbursement formula. Approximately 13.6 percent of the national sample falls into this uncertified category. Should the cost function used to determine the reimbursement rate be based on the operating financial characteristics of homes which are not eligible for reimbursement? Are certified and

non-certified homes intrinsically similar in terms of the costs of care via differences in the type of patient served, staffing levels or service intensity? Since rate setters in general answer both questions negatively, such homes were eliminated from the sample.[1]

In 26 homes the number of Medicare and Medicaid patients in SNF or ICF beds was reported such that the sum of patients in these three categories was greater than the reported census. In five of these cases the sum is one patient more than the reported census. In two other cases, it appears all patients were misclassified as Medicare since the number of patients classified as Medicaid in either SNF or ICF beds exactly equals census. The patients were treated as Medicaid. In another case all beds were certified as both SNF and ICF, and all beds were occupied by Medicaid patients. Exact double-counting of Medicaid patients occurred in this case. For the remaining 16 homes, several classification schemes are compatible with the inconsistent data.

Forty-three homes did not fill out the resident questionnaire, hence we have no data relating to patient mix. These homes were dropped from the sample.

One county reported a population density of zero. This was replaced by the average of those nonzero values for population density in the corresponding state.

There were eight cases of the input price variable equal to zero. Whenever a zero value was reported, it was approximated as the average of those nonzero values which fell into both the same SMSA classification and the same state.

---

[1] Some rate setters may wish to consider only the costs of homes with a reasonable numer of Medicaid patients and eliminate homes for cost estimation purposes with few public patients, e.g., 10 percent. This was not done in our analyses, since the interest was in a somewhat general analysis of operating costs in nursing homes.

There were 330 homes which reported a zero value for nursing home beds. A zero value was replaced by the prediction of the model.

$$NHBEDS_i = \sum_{j=1}^{4} {}_jR_{ij} + {}_5(POP65M + POP65F)_i + U_i$$

where R_{ij} = 1 if home is in region j (j = 1, 2, ..4), 0 otherwise; and (POP65M + POP65F) is the number of people over 65 in the country to which the home belongs. Only nonzero values for NHBEDS, POP65M and POP65F were used in obtaining OLS parameter estimates.

Missing data on the regulatory variables were estimated as the global average of homes for which we have data. States and homes for which this was done are as follows:

State	GSA State Code	# Homes in State
Arizona	4	1
District of Columbia	11	2
Nevada	32	1
Rhode Island	44	5
Vermont	50	3

Where staff requirements were not applicable (179 cases), the value was set to zero.

T-3.4 Model Specification and Functional Form

In this section we discuss first, the selection of plausible mathematical forms for estimating the cost functions and second, the selection of techniques for estimation of the parameters of the cost functions. We then discuss some econometric issues which must be dealt with in the course of model estimation.

T-3.4.1 Functional Form of the Cost Function

With regard to parametric specification, it was desirable that the functional form of the cost function be capable of reflecting decreasing and increasing average costs over the sample space as the output measure and utilization measure increase. The specification should also be able to reflect a traditional U-shape over the sample space with respect to average

costs as both of these measures increase. Thus, the cost equation should be such that it is capable of reflecting either of two possible phenomena: increasing returns to scale (strictly decreasing average costs) over all levels of output examined and increasing, then decreasing returns to scale (U-shape average costs as output increases). Each of the functional forms discussed below has these properties.

The first specification which one can try is given by

$$AC = a_0 + a_1Q + a_2Q^2 + b_1X + b_2X^2 + c_1Z_1 + c_2Z_2 + \ldots + c_nZ_n + U$$

where AC = average operating costs, Q = units of output supplied (i.e., average daily census or bed size, X = occupancy rate, and Z_1, Z_2, . . ., Z_n comprise a set of additional explanatory variables (e.g., case mix, urban/rural status, etc.), and U = the residual term. This form of the cost function is capable of reflecting decreasing average costs as well as the U-shape with regard to both the output measure and the occuancy rate.

A second specification is given by

$$\log(AC) = a_0 + a_1Q + a_2\log(Q) + b_1X + b_2\log(X) + c_1Z_1 + c_2Z_2 + \ldots + c_nZ_n + U$$

where log(AC) denotes the natural logarithm of the variable in brackets and AC, Q, X, and Z_1's, and U are defined as before.

A third related specification which should be explored is given by

$$\log(AC) = a_0 + a_1Q + a_2\log(Q) + b_1X + b_2\log(X) + c_1Z_1 + c_2Z_2 + \ldots + c_nZ_n + U$$

where all terms are defined as before. The second and third specifications are capable of exhibiting both decreasing costs with respect to Q and X as well as the traditional U-shape. Note that each of the three mathematical forms as written here are linear in the parameters (i.e., the a's, b's, and c's). The estimation process is considerably simpler for models which have this property.

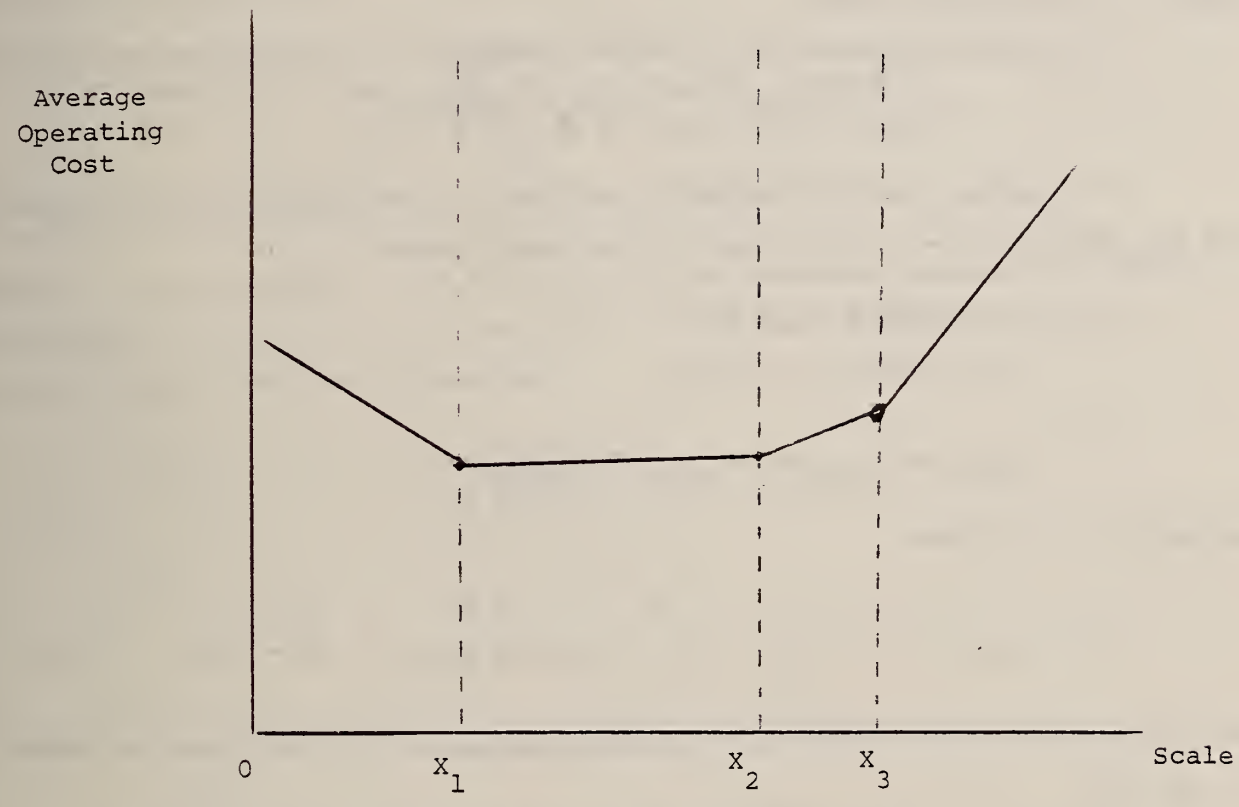
Another specification which is perfectly general is one which allows for a piece-wise linear approximation to an arbitray, but unknown functional form with regard to key variables. This is accomplished by estimating separate linear segments over different ranges of the explanatory variable as shown in Figure T-3.1. For example, if the explanatory variable in that figure were a scale factor and the dependent variable, average operating expenses, the results would indicate economies of scale between 0 and X_1, no economies between X_1 and X_2, some diseconomies between X_2 and X_3 and substantial diseconomies above X_3. Throughout the estimation work reported in Technical Papers #4 through #7, linear segments of the scale measure and the occupancy rate were "grafted" together in this way. The reasons for choosing this parametric form over the others discussed above are:

- the exploration of the existence of economies of scale and utilization in different regions of the production space is achieved;
- the degree of multi-collinearity in the data matrix is substantially reduced when compared with the usual techniques of fitting a quadratic in either of these measures;
- the response surface is continuous in these two variables;
- the response surface is easy to estimate (i.e., linear in the parameters); and
- the computation of optima is simple.

Initially in the national analyses, ten linear segments were grafted in both the output measure and the occupancy rate. This allowed for determination of a set of joint points sufficient to produce a similar approximation with a smaller number of grafted segments. By estimating several equations, each with different sets of linear segments, it was possible to observe the results converge on the most robust set of segments. The derivation of the grafting technique and the appropriate columns of the input data matrix when N linear segments in a variable, say Q, are to be grafted can be shown as follows:

Figure T-3.1

An Example of a Piece-Wise Grafting Approach



Suppose that N linear segments in Q will be grafted at the N-1 join points $K_1, K_2, \ldots, K_{N-1}$. Define variables $Q_1, Q_2, \ldots, Q_N$, and $I_1, I_2, \ldots I_N$ such that:

$$Q_i = Q \quad \text{for } K_{i-1} \leq Q < K_i$$
$$= 0 \quad \text{otherwise}$$

and

$$I_i = 1 \quad \text{for } K_{i-1} \leq Q < K_i$$
$$= 0 \quad \text{otherwise}$$

where K_o is defined as zero.

A model containing the N linear segments in Q can now be written as

$$(1) \qquad AC = \sum_{i=1}^{N} (\beta_{2i-1} I_i + \beta_{2i} Q_i) + \sum_{i=1}^{M} \gamma_i Z_i$$

To insure that the response surface is continuous we must incorporate the following N-1 constraints in the model given in (1),

$$\beta_1 + \beta_2 K_1 = \beta_3 + \beta_4 K_1$$
$$\beta_3 + \beta_4 K_2 = \beta_5 + \beta_6 K_2$$

$$(2) \qquad \beta_{2N-3} + \beta_{2N-2} K_N = \beta_{2N-1} + \beta_{2N} K_N$$

Rewriting (2), we have

$$(3) \qquad \beta_{2j+1} = \beta_1 + \sum_{i=1}^{j} \beta_{2i} (K_i - K_{i-1}) - \beta_{2j+2} K_j \quad (j = 1, 2, \ldots, N-1)$$

Upon incorporating the continuity constraints given in (3) into the model (1), we get

$$(4) \qquad AC = \beta_1(1) + \sum_{i=1}^{N-1} \beta_{2i} (Q_i - I_i K_{i-1} + (K_i - K_{i-1}) (\sum_{j=i}^{N-1} I_{j+2})) + \beta_{2N} (Q_N - I_N K_{N-1}) + \sum_{i=1}^{M} \gamma_i Z_i$$

or

$$(5) \qquad AC = \beta_1 + \sum_{i=1}^{N} \beta_{2i} X_i + \sum_{i=1}^{M} \gamma_i Z_i$$

where $X_i = 0$ for $Q \leq K_{i-1}$ $\quad (i=1, 2, \ldots, M-1)$

$$= Q - K_{i-1} \quad \text{for } K_{i-1} \leq Q < K_i$$
$$= K_i - K_{i-1} \quad \text{for } K_i \leq Q$$

and $X_N = 0$ for $Q < K_{N-1}$

$$= Q - K_{N-1} \quad \text{for } K_{N-1} < Q$$

By entering X_1, X_2,..., and X_N in the data matrix, the estimated coefficients will be the estimates of β_2, β_4, ... and β_{2N} rspectively, and the constant term in the model will be estimating β_1. Estimates of the intercept terms, β_3, β_5, ... through β_{2N-1} can be obtained via solution to the system given in (2) or (3).

An example will illuminate the technique. Suppose we desire to graft 3 linear segments in the occupancy rate using 60 and 90 as the two join points. Letting Q equal the occupancy rate, we can write our cost model as given in (1) i.e., as

$$\begin{aligned} OPEXP = {} & \beta_1 + \beta_2 Q + \Sigma \gamma_i Z_i + U \quad \text{for } 0 \leq Q < 60 \\ & \beta_3 + \beta_4 Q + \Sigma \gamma_i Z_i + U \quad \text{for } 60 \leq Q < 90 \\ & \beta_5 + \beta_6 Q + \Sigma \gamma_i Z_i + U \quad \text{for } 90 \leq Q \end{aligned}$$

where U is a disturbance term and the Z_i's are other explanatory variables. In addition to the Z_i's and a column of 1's, we must enter the following three variables into the regression:

$$\begin{aligned} X_1 &= Q && \text{for } 0 \leq Q < 60 \\ &= 60 && \text{for } 60 \leq Q \\ X_2 &= 0 && \text{for } Q < 60 \\ &= Q - 60 && \text{for } 60 \leq Q < 60 \\ &= 90 - 60 && \text{for } 90 \leq Q \\ X_3 &= 0 && \text{for } Q \leq 90 \\ &= Q - 90 && \text{for } 90 \leq Q \end{aligned}$$

The estimated coefficients of X_1, X_2, and X_3 will be $\hat{\beta}_2$, $\hat{\beta}_4$, and $\hat{\beta}_6$ respectively. The estimated intercept term will be $\hat{\beta}_1$. If upon estimation we obtained $\hat{\beta}_1 = 10$, $\hat{\beta}_2 = 2$, $\hat{\beta}_4 = -1$, and $\hat{\beta}_6 = 1$, then by solving the system given in (2), i.e.,

$$\begin{aligned} 10 + (-2)(60) &= \beta_3 + (-1)(60) \\ \beta_3 + (-1)(60) &= \beta_5 + (1)(90) \end{aligned}$$

we obtain estimates $\hat{\beta}_3 = 40$ and $\hat{\beta}_5 = -110$.

In the Technical Papers reporting the empirical results, we plot out the estimated curve with respect to both patient-days provided (or beds if used as the scale measure) and the occupancy rate.

The second stage of the econometric cost analysis for each study involved choosing one of the four parametric forms of the cost function presented for further investigation and refinement.

In the national study, for example, we initially fit all four models to the data by ordinary least squares (OLS) regression analysis.[1] Using our best judgement, it appeared that the grafted specifications were appropriate for further refinement. Although somewhat more complicated than the other forms, it gave a superior fit and alleviated multi-collinearity problems present with the other approaches. Further, it allowed the data to define its own shape, rather than to use one of the other techniques which constrain it to a specified functional form.

T-3.4.2 Methodological Issues

There are several general methodological issues relating to model specification. They are discussed in turn below.

There are certain variables that have been included in cost function estimations for nursing homes that are legitimate causes of price variation. The question is whether nursing homes should be legitimately compared for cost control purposes based on these variables. The question is really of a policy nature, and must be determined by the particular agency in question. However, there are cases where strong arguments can be made for exclusion of certain variables in the final specification of the cost equation. Some examples include the occupancy rate and factor prices.

While certain input factors in the home are quasi-fixed in the short run (e.g., nursing staff or maintenance personnel), it is unclear to what extent homes should be reimbursed for low occupancy and unused beds. Rate setters may want to consider allowing for a certain amount of slack, but also consider the establishment of an occupancy cutoff point (e.g., 90 percent) below which empty beds will not be reimbursed. This occupancy penalty may be varied for say, certain rural homes where low occupancy may be

---

[1] Ordinary least squares (OLS) estimation techniques are a special type of regression analysis in which very basic assumptions are made concerning the residual term. It is assumed that the residual terms corresponding to different homes are uncorrelated with one another. Also, it is assumed that, on average, the disturbance is zero for all homes, and that the variance associated with the disturbance is the same across all homes.

preferable to eliminate problems with accessibility. An appropriate value could be established with the aid of local planning agencies. At any rate, a policy decision must be made, and incorporated into the method used for reimbursement. In the analyses reported in Technical Papers #4 through #7 we consistently found no occupancy impact on costs. However, the variables were left in the model to present rate setters with evidence that such variables do not (in our data) affect operating costs. If significant economies of utilization, however, had been observed, the inclusion of this variable would have been an issue since its presense would imply legitimately high costs for low occupancy rates.

A second area of concern is factor prices, particularly wages. The issue here is the extent to which nursing homes have control over wage rates. That is, if nursing home operators have control over wage rates (or purchase vs. lease arrangements for capital equipment and supplies), then including factor prices in the cost equation as legitimate reimbursable costs may simply allow for a cost pass-through of what are really controllable items. One possible solution to this problem is the use of market-area average wage data for relevant occupations instead of facility-specific data. While an operator may be able to inflate reported employee wages in an effort to receive a higher reimbursement rate, he (she) has much less control over market-area wages. One can argue, in addition, that the rate setter should use of set of wages exogenous to the control of the nursing home industry. The culmination of local nursing homes together exerts considerable influence over local levels of nursing wages, hence wages for a broader occupational group would be appropriate (e.g., a group of service industry workers).

T-3.5 Model Estimation

Throughout our work, cost equations were estimated by ordinary least squares which are a special case of regression analysis. The estimators associated with ordinary least squares (OLS) can be shown to have optimal properties when estimating the parameters of a correctly specified model. It is generally assumed that the residual term, U, has an expected value of zero, the variation associated with the residual term is constant across all observations, and residuals associated with different observations are uncorrelated. Moreover, it is assumed that no linear relationships exist

between explanatory variables in the model, and that each explanatory variable is uncorrelated with the residual term.

Below we address, in turn, several econometric issues relevant to the assumptions of OLS in cost function estimation.

T-3.5.1 Residual Behavior

After obtaining an estimated departmental cost function, the residuals should be examined in detail. Special concerns which should be watched out for include:

(1) The estimated residuals of nursing homes of a certain type (e.g., large bed size or rural location) which exhibit larger or smaller dispersions than estimated residuals of homes not of that type (e.g., small bed size or urban location).

(2) A strong positive correlation existing between the estimated residuals of nursing homes of a certain type (e.g., low occupancy rate, extensive scope of services, etc.)

(3) Both larger or smaller dispersions and positive correlations existing between the estimated residual terms of nursing homes having a certain characteristic when the characteristic was unaccounted for in the cost functions.

The first situation suggests that the assumption of constant variance across the residuals of all observations was violated. When this assumption is invalid, residual terms are said to be heteroskedastic. The second situation suggests that the assumption regarding the independence of residuals associated with different facilities was violated. If the assumption is invalid, residuals are correlated. The third situation, a combination of the first and second, suggests that the residual terms were both correlated and heteroskedastic.

Each of the three situations calls for the application of generalized least squares (GLS) regression techniques instead of OLS procedures. GLS regression techniques differ from OLS procedures in terms of the assumptions made regarding the residual. The detailed statistical properties of OLS and GLS estimators and the consequences of violating the assumptions inherent in estimation of the model are well-known and are reported elsewhere.[1]

---

[1] The statistical properties of OLS and GLS estimators, the consequences of violating assumptions in either model, the mathematical formulation of each of the three GLS models, and the appropriate estimation procedures for all models are discussed in Kmenta (1971), Pindyck and Rubinfeld (1976), and Johnston (1972).

Cost equation residuals should be examined carefully to determine if any of the three situations mentioned above exist. A detailed residual analysis might have the following components:

(1) Examination of the characteristics of facilities with very large and very small residuals to detect any patterns. This can be done by examining the means and standard errors of variables such as operating expenses, average daily census, occupancy rate, etc. for homes with residuals in various ranges. This will reveal potential model bias, i.e., whether the model consistently over-predicts or under-predicts homes of certain type.

(2) Examination of the mean and standard error of the disturbance for residual subgroups based on urban-rural status, certification status, ownership types, regional location, etc. This will uncover any problems of heteroskedasticity.

A thorough residual analysis was done for the national analysis and found that the possible problems enumerated above were not present. (See Technical Paper #4.)

T-3.5.2 Multi-collinearity

It was probable that multi-collinearity among the predictors would be a problem. Multi-collinearity would not affect the aggregate predictive performance of the cost equations, assuming one did not extrapolate beyond the range of prediction. Hence, it would not affect residual analyses designed to identify outlying nursing homes. However, multi-collinearity would affect the reliability of the model for explanatory purposes since it affects the estimated distribution of regression coefficients. Specifically, the main consequences of multi-collinearity are that certain parameter estimates are unstable, the estimated variance of specific parameter estimates is biased upwards, and parameter estimates may be highly correlated. By instability we mean that the estimated coefficients become very sensitive to the sample size insofar as the addition or deletion of a few observations can sometimes change estimates drastically.[1] An examination of regression coefficients in terms of sign, magnitude, and reliability, and an evaluation

[1] Mason et al. (1975) and Fisher (1977) have argued that in addition to these effects, estimated coefficients will generally be large in absolute value when multi-collinearity is present.

of each independent variable in terms of its contribution to explaining variation in operating costs were two important and intersting areas of analysis. In each analysis explanatory variables were examined for near collinearities. In most cases, the final versions of cost equations reported here are free from an excessively high degree of multi-collinearity. The main problem in our analyses involved the numerous patient descriptors. In general, it appeared that when a patient was "sick," he was "debilitated" along a wide variety of measures of activities of daily living (ADL) and diagnoses in consequence, while the sum of the estimated coefficients of the patient characteristics was unbiased, the coefficients on indiviudal ADL and diagnoses were probably biased. In an effort to reduce multi-collinearity, we reestimated the regression, in a somewhat arbitrary way by eliminating those variables whose "t statistic" was less than one, i.e., the estimated coefficient was less than the standard error. The main consequence was to eliminate most of the patient descriptors; the coefficients on the remaining variables remained much as in the complete mode and the summary statistics remained virtually the same. In brief, we eliminated much of the multi-collinearity at no loss of accuracy to the rest of the variables.

T-3.6 Using Cost Functions in a Rate-Setting Environment

It is helpful to recognize the strengths and weaknesses associated with the regression approach to cost function estimation. It is also important to consider how the estimated functions are useful to rate setters. The next two sections discuss the role of the cost equations in a rate-setting environment.

T-3.6.1 What Can Rate Setters Learn from Cost Equations?

Specifically, an analysis of nursing home cost equations provides the following:

(1) A list of nursing homes which appear as high cost.

(2) A list of facilities that appear to be low cost outliers. (The usefulness of this information is discussed below.)

(3) Information suggesting the relative effect of independent variables on unit costs. Also included is information on the ability of the independent variables to explain costs.

(4) Information on the approximate dollar value bonus of penalty associated with facility characteristics such as certification status, location and sponsorship of the facility.

(5) A list of the types of data that are useful in explaining cost variation and which should be gathered by rate setters.

(6) An equitable method of identifying high cost nursing homes in the sense that the residuals of all facilities in the sample are given equal weight in the course of model estimation.

Each of the six types of information is useful in a budget review or rate-setting context. They are discussed in turn below.

- Identifying High Cost Outliers

The first, an agenda containing a list of nursing homes that should be examined more closely, can be generated in a fairly systematic fashion by isolating those facilities which have large residuals on the cost equations. Although not done in this study, such an agenda could be extremely useful in any budget review system. The regression analysis approach attempts to quantitatively account for those factors which logically cause costs to vary. The approach is relatively simple to use and most likely less time-consuming than a detailed nursing-home-by-nursing home review. One clear advantage of the estimated cost function over other approaches to outlier identification is that a large amount of information which is useful to rate setters emerges via the estimation process. By estimating cost functions, as was done here, one comes closer to determining which institutional or case-mix features have a systematic and measurable impact on costs. This windfall of information can be used to evaluate the reasonableness of the facility budget and the reimbursement rate structure. However, there are some weaknesses with the approach which we identify in the next section.

- Identifying Low Cost Outliers

Similarly, the estimated cost equations provide a means of identifying providers which are relatively <u>low cost</u> given their facility characteristics. And so, the equations can be used as <u>ad hoc</u> tools to detect a group of nursing homes which appear as economically efficient on different cost centers. An examination of homes in this low cost category may provide insights on efficient production schemes which is valuable information in any rate-setting system, especially if one desires to set guideline "efficient" production methods. A screening of low cost outliers may also uncover possible deficiencies in the services provided by some facilities.

Finally, with regard to the types of useful information generated through the residual analyses, rate setters may discover quirks in the accounting procedures of individual facilities. Ideally, accounting cost definitions should be the same across nursing homes. Uniform definitions facilitate the interpretation of the empirical results and provide the basis for meaningful comparisons of cost behavior across facilities. It is highly likely that the interpretation of some definitions used in the various Medicaid cost reports may not be consistent among reporters across different homes. Outlier analysis is one tool which helps uncover accounting inconsistencies. Unfortunately, until uniform accounting and reporting are truly in place, sophisticated methodologies may be of little help in calculating the cost of a specified product for rate-setting purposes. The development of these methodologies, however, should continue nevertheless in the hope that consistent and reliable data will be forthcoming.

- Measuring the Relative Impact of Provider Characteristics

The estimated cost equations also serve the function of identifying for rate setters those factors which have the most "leverage" in affecting costs. Leverage can be assessed through the examination of beta coefficients. Beta coefficients are a by-product of the estimation procedure. These standardized coefficients adjust the estimated slope parameters by the ratio of the standard deviation of the independent variable to the standard deviation of unit costs. Thus, the beta coefficients are somewhat comparable to one another and indicate the relative impcat of each predictor. (A beta coefficient of 0.4 can be interpreted to mean that a 1 standard deviation change in the independent variable will lead to a 0.4 standard deviation change in unit costs.) Thus, the cost function coefficients can provide a base of information about cost impacts of the variables. This type of information is useful to planning agencies in that it facilitates the assessment of ways to influence the movement of costs into target or goal ranges set by planners.

- Measuring the Absolute Impact Provider Characteristics

Closely related to information on the relative effects of nursing characteristics (the beta coefficients) is information providing an approximate estimate of the absolute or dollar value impact of various predictors. This information is contained in the regression slope coefficients. Each coefficient

provides a rate setter with a dollar value estimate of the marginal effect of changing the value of the predictor. Thus, rate setters can assess the marginal impact of various facility features. For instance, it allows assessment of the marginal cost of increasing output. Also in approving nursing home budgets or in setting nursing home reimbursement rates, information on regression coefficients can be used by rate setters as a tool in assessing the size of the average dollar bonus or dollar penalty provided for different institutional features of the home.

- Indicating Needed Data

Finally, by estimating cost functions rate setters obtain information on the types of data which can and cannot explain cost variation, thus indicating where to focus future data collection efforts.

- Equitability Across the Sample

A final beneficial feature of the regression cost equation is that they are equitable across providers in the sense that the estimation error associated with each observation in the sample is given equal weight in deriving the estimated parameters of the equation. This is a basic property of OLS regression which results in "equitability" in computing the cost of a specified product.

We turn now to some of the inherent difficulties with the proposed methodology.

T-3.6.2 A Problem with the Regression Approach Outlier Identification

In this section we discuss a methodological issue which serves to caution users of the methodology on the limitation of the regression approach to cost function estimation. Briefly, the issue is:

> Regression analysis is unable to detect economic efficiency or inefficiency in selected facilities, and as such may need to be supplemented by other screening tools.

It should be noted that least squares techniques are not the only approach to estimating cost functions. Envelope programming procedures are an alternative approach. Although less traditional, the latter procedures allow one to compute a cost curve which gives more weight to those nursing homes which are "best practice" in the sense of having relatively small per unit costs for each possible output. The application of envelope programming

estimation procedures is compatible with the premise that facilities which have the least costs for each unit output are the most efficient. Our approach, however, in estimating the cost functions was to give equal weights to all homes in the sample.

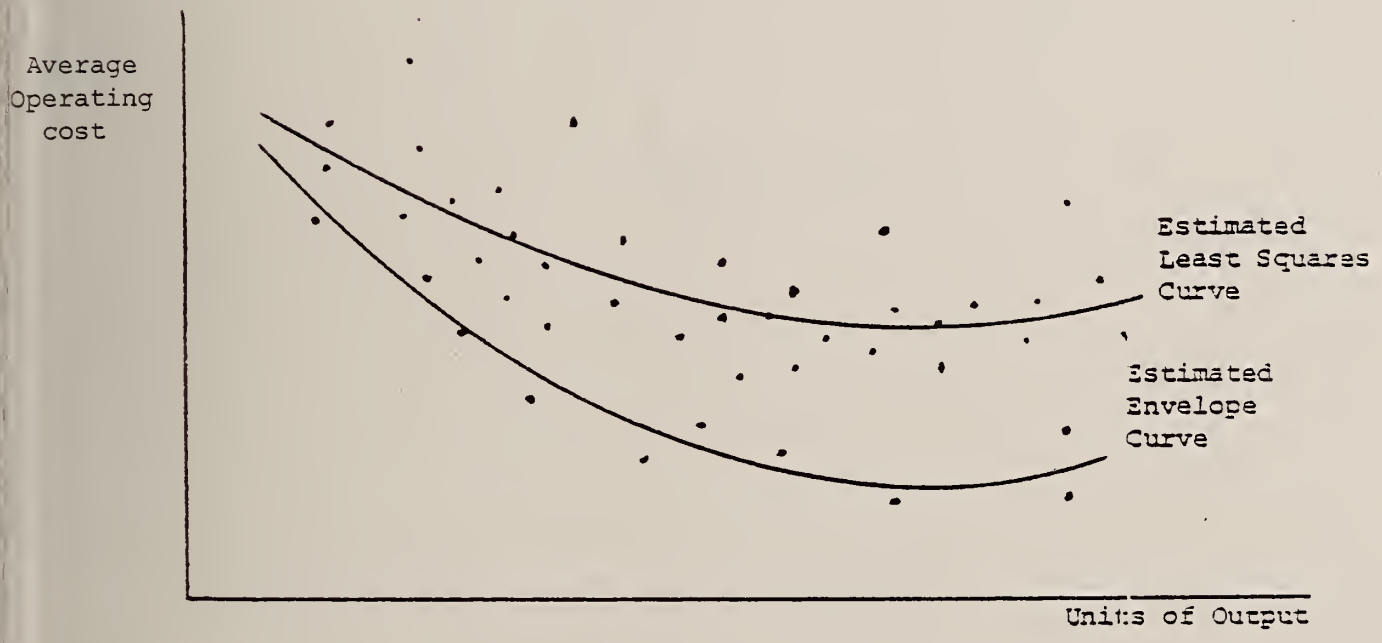
Graphically, in two dimensions, it is easy to note the distinction between the two approaches to fitting a cost model: least squares techniques and an approach such as envelope programming procedures. Consider Figure T-3.2 below. Each facility in our sample is represented by a point with coordinates of average operating costs and output supplied. Estimating a least squares curve we gave equal weight to minimizing the estimated residuals associated with all facilities. The result is that the estimated cost curve predicts, in a sense, "average costs" for each unit of output. In estimating the envelope cost curve, however, we give more weight to minimizing the estimated residuals associated with homes which have the least costs for different levels of output. The result is that the curve fits best for these least-cost nursing homes. By letting the envelope curve set the standard for "efficient" cost behavior, far more nursing homes appear as inefficient.

Another alternative to the least squares technique applied here would be to fit a curve to a selected sample of facilities which were viewed as best practice and which represented a wide range of alternative types of nursing homes. This approach is the most subjective of the three discussed here since the estimation sample is selected. However, it has the advantage of allowing users to incorporate their own efficiency standards into the cost function analysis.

A point of contrast between this and the OLS technique is that in using the selected sample approach, situations are avoided where an equation is unable to detect inefficiency because it was fit to a sample of homes which were generally inefficient. Using OLS, an equation may have an R^2 of 0.75 and reveal no high cost outliers, even though an objective case-by-case study would identify many facilities as inefficient.

Figure T-3.2

Graphic Distinction Between an OLS and Envelope Curve Estimation



TECHNICAL PAPER #4

Technical Paper #4

AN ANALYSIS OF NURSING HOME COSTS BASED ON NATIONAL SAMPLE*

T-4.0 Introduction

Ever since the Medicaid and Medicare programs began to pay directly for nursing home care, increasing public and governmental attention has been paid to the nursing home industry. National expenditures on nursing home care have risen from $1.7 billion in 1967 to $10.6 billion in 1976. In "real" dollar terms, nursing home expenditures grew by around 370 percent.[1] In response to rising expenditures, scandals, and concern over inappropriate patient, the federal and state governments have proposed and implemented several rounds of regulations. The most recent effort with regard to reimbursements at the federal level has been Public Law 92-603, which mandates that state Medicaid plans shall provide for payment of skilled and intermediate care facilities on a "reasonable cost-related basis." Determination of the parameters of a reimbursement mechanism requires an understanding of the many factors which affect nursing home costs, including input prices, output characteristics, patient characteristics and the presence of regulations themselves. Such information is necessary for the design of a reimbursement system which will structure a set of incentives to lead nursing home operators to make rational management decisions which in turn leads to the fulfillment of public policy goals. A socially desirable reimbursement system should meet this, as well as other criteria.

The objective of the larger study, of which this paper is a part, is the development of a prototype formula prospective reimbursement[2] system

---

*by Gail Jensen and Howard Birnbaum.

[1] The consumer price index was 100 in 1967 and 170.5 in 1976. If the deflation factor is taken as the medical care component of the consumer price index, the "real" growth in nursing home expenditures was around 340 percent; the medical care price index was 100 in 1967 and 184.7 in 1976 (U.S. Census, 1977).

[2] Dowling (1976) defines prospective reimbursement or rate setting as a "cost containment strategy wherein an external authority establishes the prices that providers are allowed to charge and/or that third parties are required to pay for specified services in advance."

for nursing home operating costs. A series of other papers produced by this study[1] have reviewed the behavior of nursing homes industry, discussed a variety of issues involved in designing a reimbursement system, and developed a research design for analyzing the determinants of variation in the operating costs of nursing homes. A major conclusion of these papers is that the design of a reimbursement mechanism must consider more than the direct determinants of nursing home costs. A thorough study must determine how providers react to reimbursement regulation in ways which affect cost indirectly, for reimbursement practices provide incentives to alter provider behavior. A single equation model which shows the relationship of input prices and output characteristics to average cost is insufficient for establishing a reimbursement system; variables treated as characteristics of firm behavior, such as the public-private mix of patients, also are determined in response to policy directives. Since providers are able to alter their behavior with respect to services offered, public-private mix, etc., in response to reimbursement and regulatory practices, cost analyses of the nursing home industry need to be embedded in a system of equations encompassing those variables under the control of nursing home operators. Knowledge of the system relationships is necessary if the reimbursement mechanism is to anticipate provider reactions and channel provider behavior in ways which meet policy goals. The overall study of which this paper is a part develops this information via a series of analyses: a cost analysis of data from a national cross-sectional sample of nursing homes, a series of cost analysis for each of three states with data for each of three years, and a multi-equation model of the behavior of the nursing home industry, using a national cross-section sample of nursing homes.

It is the purpose of this paper to report the econometric results of the cost analysis of a national sample of nursing homes. Later papers will integrate the results of the national cost analysis with the state-specific cost analyses and with a multi-equation model of the nursing home industry. These analyses will then form the basis for development of a prototype prospective reimbursement system.

---

[1] Bishop, Reagan, and Birnbaum, 1977; Birnbaum, Bishop, Jensen, Lee, Reagan, Wilson, and Zook, 1977; Bishop and Birnbaum, 1978.

The remainder of this chapter is organized as follows. Section T-4.1 reviews the conceptual model. Section T-4.2 discusses empirical issues including model specification, variables and the data used here. Section T-4.3 presents results from the complete specification of the model. Because the complete model includes a large number of variables and allows the potential for multi-collinearity, Section T-4.4 reviews results from a "limited" model that includes a restricted list of explanatory variables. Section T-4.5 summarizes the results. Appendix A presents additional information on the data and supplemental tables. Appendix T-4.B presents an analysis of the determinants of cost for individual cost centers, e.g., wages, laundry, food, etc.

T-4.1 Review of Theoretical Development

This section briefly develops a theoretical approach to estimating nursing home operating cost functions. The cost function of economic theory is a schedule relating the minimum costs of production to the output produced and the prices of inputs (Walters, 1963). As output and input prices increase, so will costs. In addition, economies of scale and utilization are possible, such that there may be decreasing average costs as the output measure and the occupany rate increase. Thus, the model should allow for the possibility of scale effects. However, the characteristics of the nursing home industry are such that traditional cost function theory requires modification for use here.

The model developed below relates the average operating costs of nursing homes (AOC) to four types of variables: vectors of output measures (Q), input prices (P), regulatory variables (R) and facility characteristics describing the production process (F). In equation form:

$$AOC = f(Q,P,R,F)$$

T-4.1.1 Output Variables

Skinner and Yett (1970) have shown for nursing homes that case flow measures of output affect costs. Because increased case flow involves additional administrative work, homes with greater patient turnover, shorter lengths of stay, and greater admissions rates are likely to have greater costs. Case flow variables also can be viewed as describing patient care provided; a short-stay patient may be more likely to require intensive and rehabilitative care than a long-stay patient using more maintenance services.

However, unlike many other industries, output in the nursing home context is not a unidimensional quantity. While the basic unit of analysis for nursing homes is typically taken to be the patient day, it is more appropriate to consider multiple measures of output. The number of patient days provided must be augmented to reflect the "quality" of output as well as patient characteristics, specifically the health status and debility of a patient. For example, a mix of patients primarily unable to feed themselves may require more nursing hours per patient day, with commensurately greater costs, than a mix of relatively self-sufficient patients. Hence, indicators of patient status (i.e., case mix) should be included in the cost model. Similarly, costs may vary by the quality of the output produced. The difficulties in measuring quality in health care research are well known. To the extent possible, quality is taken into account by incorporating service variables in the model, such as the presence or absence of occupational therapy, the percentage of one-bed rooms, etc.[1] Very little quality-specific information was available in the data used here (or in other nursing home data); its absence and the use of services as proxy variables raise issues of interpretation since one cannot distinguish between services which are inappropriate to patients or inefficiencies. In any event, the output of a nursing home is a vector of characteristics including patient days, as well as case mix, quality, and services.

T-4.1.2 Regulation Variables

Because the nursing home industry is regulated, a second set of modifications to the traditional cost functions of economic theory are required. In an unregulated industry, facility type need not necessarily be distinguished to the extent that relevant production differences are captured by the patient and service characteristics. However, since states typically impose differential requirements (e.g., nurse staffing requirements) and may explicitly consider facility type in establishing reimbursement rates, differential regulatory requirements suggest extraordinary cost differences for two reasons. First, as a consequence of regulation, historical cost differences may be maintained over time, independent of patient status, "quality" of care, etc. Second, regulations may impose cost differences because of differential resource requirements, as for example by imposing

---

[1] We discuss elsewhere (Birnbaum, et al., 1977) how the union of information on patient mix and services provide a rough approximation of "quality."

staffing requirments. Therefore, it is necessary to take regulation and reimbursement practices explicitly into account.

T-4.1.3 Other Explanatory Variables: Input Prices and Facility Characteristics Describing Production Techniques

In addition to output and regulatory variables, the cost analysis must consider two other types of variables. The first includes measures of input prices. The last vector of variables which may affect nursing home characteristics includes an *ad hoc* set of facility characteristics. The facility characteristics reflect alternative methods for producing comparable output, and may reflect, for example, managerial decisions and abilities to operate on the general cost curve facing all nursing homes.

T-4.2 The Empirical Model

This section develops the empirical specifications for estimating nursing home operating cost functions along the lines discussed above. Section T-4.2.1 discusses the data used in the analysis; Section T-4.2.2 discusses variable specification and hypothesizes expected relationships for the explanatory variables. Section T-4.2.3 discusses selection of suitable techniques for estimation of the parameters of the cost function.

T-4.2.1 Data

Data used in these analyses are from the the 1973-74 National Nursing Home Survey conducted by the Division of Health Resources Utilization Statistics of the National Center for Health Statistics (NCHS). This nationwide survey provided information on the sampled nursing homes, their residents, and staff. Of the 2,118 facilities sampled by the NCHS, 525 were non-certified homes and were eliminated from the set of observations used in the empirical work. Another 470 homes were deleted because of lack of data on the dependent variable: average operating costs, or other variables.[1] A

[1] Average operating cost was formed as the sum of facility costs for all payroll and fringe benefits, food, drugs, supplies, maintenance, services for contract, utilities and miscellaneous expenses. Observations lacking one or more of these data were deleted. Also deleted were homes with severe discrepancies between annual patient days and the quantity (bed size times 365), homes with extremely high values for admissions or discharges (i.e., greater than 1500), or homes with an extremely low census or bed size (i.e., less than 4). Facilities which were extended care units of hospitals too were deleted (since hospital-based facilities might be specializing in shorter-term care and thus might not be comparable to freestanding clinics). Also deleted were facilities which did not fill out the resident questionnaire (since the resident questionnaire was the source of patient specific data). Finally, facilities in those few (and small) states for which reimbursement information was not available (i.e., Arizona, Alaska, and D.C.) were deleted.

total of 1,125 nursing homes were used in the cost analyses. All data on costs incurred by the facility as well as some facility characteristics were collected for calendar year 1972, but if the home's accounting records were on another basis (e.g., fiscal year) the data apply to the most recent fiscal year prior to the survey. Some of the data do not apply to the 1972 calendar year; rather, they apply to the period in 1973-74 during which the survey was conducted. In addition to these time period differences, the reader is referred to Appendix T-4.A which contains further information on these data.

In addition to the NCHS data, the study utilizes data from the Area Resource File (ARF) characteristics of the county in which the nursing home is located.

There is no single, comprehensive source of regulatory/reimbursement data. As a result, data concerning the regulation of nursing homes in 1972 were developed from a variety of sources including surveys conducted by the Urban Institute (UI), Applied Management Sciences (AMS), Battelle Health Care Study Center, the Government Accounting Office (GAO), and the DHEW Social and Rehabilitative Services (SRS). The data were verified and modified, if necessary, with information on state-specific regulatory practices found in the other data bases. Table T-4.1 presents these variables. Appendix T-4.A discusses these variables, data sources, and procedures for reconciling variations. The result was a unified set of regulatory/reimbursement variables.[1]

---

[1] For empirical purposes, a "prospective" reimbursement system refers to a rate setting mechanism in which the rate for an individual facility for a specific period is determined prior to the period in question with no automatic retrospective adjustment. A "flat rate" system refers to a rate setting mechanism in which the rate for a specific period is set at one level for a group of facilities, where the rate may be negotiated, legislated, or established via a point system. A "cost based" system refers to a rate setting mechanism in which the rate for a specific period was determined during or after the period in question; typical procedures here are to follow the Medicare guidelines paying all "allowable costs," or alternatively to pay "cost related," or "actual cost" rates.

The following decision format was used to resolve differences among the sources with regard to reimbursement mechanism. If all sources agreed that the rate setting system was a "cost-based," "prospective," or "flat/point" system, the common answer was accepted. If the UI and SRS data agreed, but

the AMS or GAO data differed, the UI/SRS information was accepted, since these later sources were for the appropriate year (GAO data were for 1970; AMS data were for 1974-75). If the UI information differed from the SRS information, we accepted the SRS data, since the data classifications were supplemented by descriptive information appropriate for evaluating the suitability of the answer. Of course, if the SRS descriptive information contradicted the categorical answer, the SRS response had to be altered as appropriate. For example, Michigan viewed its system (p. 5 of questionnaire) as a "cost system;" however, in reviewing their comments, it became clear that it was, in fact, a prospective system, under a definition of "prospective" as a rate setting mechanism in which the rate for an individual facility for a specific period is determined prior to the period in question with no automatic retrospective adjustement. A final check was to review inconsistencies with other researchers using these same sources of data.

Table T-4.1

Sources of Reimbursement and Regulatory Variables

Variable[1]	Source[2]				
	AMS	UI	GAO	SRS	Battelle
Payment Mechanism: prospective; flat rate; cost-based	X	X	X	X	X
Cost Basis: class of facilities or individual	X				
Payment Limits	X				
Occupancy Limits Present	X				
Uniform Chart of Accounts Required	X				
Life Safety Codes in Effect		X			
Medical Assessments Required Prior to Admission		X			
Certificate of Need in Effect		X			
Level of Staff Requirements		X			

[1]Definitions of these variables are as follows:

Payment Mechanism: Prospective rate--rate for an individual facility for a specific period is determined prior to the period in question with no automatic retrospective adjustment; Flat rate--rate for a specified period is set at one level for a group of facilities, and may be negotiated, legislated, or established by a point system; Cost based rate--rate for a specified price determined during or after the period in question; typical procedures here are to follow the Medicare guidelines of paying "allowable costs" or alternatively to pay "cost related," or "actual cost" rates, etc.; Cost Basis--describes whether rates are established for a facility in reference to a class of facilities or on an individual basis.

Payment Limits: Describes whether the state specifies either limits on total per patient payments per unit of time and/or limits on specified categories of costs.

[2]AMS refers to the Applied Management Sciences Study (1977); UI refers to the Urban Institute Survey (N.D.); GAO refers to the Government Accounting Office Study (1972); SRS refers to the DHEW Social and Rehabilitative Services Survey (1973); Battelle refers to the Battelle Study (N.D.). Although the AMS data refers to the period 1974-75, they are the only data on these variables, and we assume that they are appropriate for 1972. UI data are for 1972.

T-4.2.2 Variable Specification

This study examines the variation in operating costs. Capital costs are not included in the analysis.[1] The dependent variable in the cost equation is Average Operating Costs (AOC) and is defined to include all payroll and fringe benefits, food, drugs, supplies, maintenance, laundry, services by contract, utilities and other miscellaneous expenses. Excluded are property and profit taxes, depreciation, interest charges on outstanding debt, rent on plant or equipment, amortization of capital improvements, and profit.[2] Details of this definition of cost must be kept in mind, since the nursing home industry is so profit-oriented. This analysis implicitly assumes that capital costs are not substitutable for operating costs in the nursing home industry. To the extent that substitution exists between capital and operating costs (for example, to circumvent rate regulations), the results are weakened; the degree of substitutability was not explored in the study.

Section T-4.1 argued that AOC is a function of output, input prices, regulation, and additional facility characteristics. In brief, the general model for the nursing home average operating cost function looks as follows:

$$AOC = f(Q, P, R, F)$$

where AOC = average (per patient day) operating costs

Q = vector of output produced by the facility, including:

- o measures of scale of production, including:
  - average daily census
  - occupancy rate

---

[1] Capital costs are difficult to measure accurately. Accounting data available here do not reflect true costs of capital (e.g., because of differences in allowing for depreciation).

[2] Profit was excluded because it may be a function of total cost or total capital of the home, or is an accounting residual.

- services
  - offered services
  - intensity of use of services
  - presence of one-, two-, and three-bed rooms
- patient mix
  - age
  - sex
  - debility measures
  - health status (i.e., diagnosis at time of survey)
- case flow
  - admissions per patient day (or length of stay)
  - percent death

P = input prices
- wage rate measures
- SMSA location

R = regulatory variables, including:
- certification (i.e., SNF versus ICF)
- region of country
- payment source (i.e., percent Medicare, Medicaid)
- specific regulations (e.g., reimbursement mechanism)

F = additional facility characteristics describing production differences
- ownership (i.e., profit/non-profit)
- chain ownership
- bed changes

The following pages briefly describe the variables used in empirical estimation and associated hypotheses. More complete definitions and sources of the variables as well as their means and standard errors are presented in the next section.

Output: Scale. It is hypothesized that the size and occupancy rate of a facility affects its costs, such that there are economies of scale and utilization. Homes producing a larger output (e.g., a larger average daily census) have lower costs; and, at a given capacity, average cost declines as occupancy rates increase. These variables also allow for differences between the short and long run cost functions. These occupancy variables (OCi) capture disequilibria deviations from the longer-run cost curve. Scale of production will be represented by average daily census (AD).[1]

Output: Services. It was argued in Technical Paper #3 that service intensity differentiates the nursing home product. Facilities providing more services can be presumed to use more resources and hence have higher costs. However, this relationship between services and "quality" is tenuous, for without explicit quality indicators, one can not distinguish whether services are appropriate for patient mix. Nevertheless, no other measures of quality are available. Thus, service variables are taken to represent "quality" differences, given a particular patient mix. It is hypothesized that the more services a facility offers, the greater its average operating cost; on the other hand, it is hypothesized that nursing homes "staff up" for a particular range of offered services appropriate to their case mix, so that the intensity of services actually provided does not affect costs, given a particular patient mix. Hence, patient mix (e.g., health status, debility) also impact cost, independent of services.

Data on the availability of certain services, whether offered or not, is present for the following seven services: (1) occupational therapy, (2) recreation therapy, (3) physical therapy, (4) patient counseling,

---

[1] AD is computed by dividing total yearly patients days by 365; OC is computed by dividing total patient days by mean operating bed size times 365. AD comes closer, in a theoretical sense, to a measure of output than a variable measuring bedsize of the facility. AD and bedsize correlate so highly (r^2 =.97), that they are virtually identical.

(5) rehabilitative services, (6) speech and hearing therapy, and (7) educational programs. These variables are noted as: (1) OT, (2) RT, (3) PT, (4) COUNSEL, (5) REHAB, (6) SPHT, and (7) EDUC respectively. In the regression, they are entered as (0, 1) dummy variables, where 1 indicates that the services is present.

Twelve variables describing the intensity of services provided are also in the model. Each of the 12 variables measures the percentage of patients in the week prior to the survey receiving one of 12 services. The services examined were: (1) nasal feeding, (2) blood pressure reading, (3) enema, (4) catheterization, (5) full bed bath, (6) bowel and bladder retraining, (7) oxygen therapy, (8) application of dressing or bandage, (9) temperature-pulse-respiration reading, (10) hypodemic injection, (11) irrigation, and (12) intravenous injection. These variables are noted as: (1) NASAL, (2) BLOOD, (3) ENEMA, (4) CATH, (5) BATH, (6) BOWEL, (7) OXYGEN, (8) BAND, (9) PULSE, (10) HYPO, (11) IRRIG, and (12) INTRAV. An alternative service intensity variable is also considered, SUMSER, which is the sum of the service units provided to patients in the week prior to the survey.

Also in the model are variables indicating the percentage of one-bed rooms (PONEB) and the percentage of three-or-more-bed rooms (PTHREEB).[1] These variables proxy the quality of services provided by plant and equipment.

In other studies, ownership and certification level have been used to measure different product types to allow for the possibility that non-profit homes treat different kinds of patients, provide different services or provide different quality of care than do for-profit homes. Also, certification level (SNF/ICF) insures that a common minimum set of services is available, though homes may provide services in excess of the minimum. Since the model explicitly includes the service variables, the ownership and certification variables are considered below as regulatory and production characteristic variables.

---

[1] By omitting the percentage of two-or-more bed rooms as a variable, the coefficients of PONEB and PTHREEB measure the differential effect on costs of one- and three-or-more bed rooms, relative to the percentage of two-bed rooms. This general procedure is used throughout the analysis, as for example, in including ICF and ICF-SNF variables, but omitting the SNF variables; there the coefficients of ICF and ICF-SNF capture the effect of certification relative to SNF status.

Output: Patient Mix Debility

All patient mix measures (with the exception of the payment source variables) were based on a sample of patients in each facility; the size of the sample ranged from 1 to 10 patients. For this reason the patient mix data represent only an approximation to an actual characterization of the case mix in each home. Each patient related variable is subject to considerable error.[1] Given this situation we can expect that the regression coefficients for the patient mix variables will be biased toward zero, i.e., it is likely that they will understate the importance of case mix in determining costs. This presents a problem in interpreting the results since we cannot reliably assess the role of case mix. However, this econometric problem is dwarfed by the absense of quality-specific data. Without quality data, there is no way to know what type of service or output is being produced. Since quality of care is related to patient mix and services provided, variations in quality mean that neither is being held constant.[2] The presense of these two problems raises questions about the validity of patient mix coefficient interpretation. Possible conclusions that the variables have little effect on costs should not be made without some notion of the size of the bias. Nonetheless, we have used the patient mix measures in our cost model. The variables used are discussed below.

Debility status measures were available in terms of six activities of daily living (ADL): (1) grooming, (2) dressing, (3) feeding, (4) ambulation to the toilet room, (5) moving in and out of a bed or chair, and (6)

---

[1] The problem is actually that of "errors-in-variables." Thus, least squares estimates of the regression parameters will be biased and inconsistent, with the degree of bias and inconsistency being related to the variance of the measurement error.

[2] Independent of the errors-in-variables issue discussed in the prior footnote, the problem of omitted quality variables introduces some bias and inconsistency into the model. We can expect that since patient mix and services are related to quality (which should properly enter the cost function) the coefficients of these variables will be both biased and inconsistent, with the degree of bias and inconsistency being a function of the correlation between quality and the patient and service mix variables.

continence. They correspond to the notation: (1) GROOM, (2) DRESS, (3) EAT, (4) WALK, (5) BED, and (6) CONTIN. Each ADL variable is defined in terms of both the percentage of patients partially dependent (noted with PD as a postscript) and the percentage of patients totally dependent (noted with D as a postscript) in the ADL category, relative to being independent. In this way, the estimated coefficient of each ADL variable can be interpreted in a straightforward manner. For example, the estimated coefficient of the percentage of patients who are, say, partially dependent in grooming (GROOMPD) is an estimate of the average increment to operating costs resulting from a unit increase (a unit being a percentage point) in the proportion of patients who are partially dependent in grooming. The costs associated with various combinations of debilities and debility levels then are measured directly.

Also investigated were other index methods of entering the debility level variables, such as the Guttman Index, the sum of scores procedures, etc., as described later in Section T-4.3.6.

Other measures of the patient mix included age, sex, diagnosis, and mental status indicators. Specifically, the following were included: the average patient age (AGE), the percentage of patients who were female (FEMALE), and the percentage of patients with the diagnosis of (a) a hardening of the arteries (ARTERY), (b) stroke (STROKE), (c) accident, poisoning, or violence (ACCIDENT), (d) endocrine, nutritional, or metabolic disease (NUTRIT), (e) senility, old age, or other ill-defined conditions (OLD), (f) respiratory disease (RESPIRAT), (g) mental disorder, e.g., mental retardation, psychoses, neuroses, mental illness, or emotional problems (MENTAL), (h) disease of the digestive system (DIGEST), (i) neoplasmic disorder (NEOPLASM), (j) heart attack (HEART). Relating to the mental status of patients, variables indicating the percentage of patients in each of the following categories were included: depressed (DEPRESS), aggitated and aggressive (NERVOUS), abusive and aggressive (ABUSE), confused or senile (SENILE), and having distrubed sleep (SLEEP).

---

[1] As pointed out in Technical Paper #3, for cost modeling purposes it may not be worth the effort to collect extensive and detailed data on a small sample of patients since only minimal gains in model accuracy are realized with the inclusion of the data. Collecting limited data on all residents would be much more useful.

Output: Case Flow. Increases in case flow increase the administrative burden of the home, and hence, it is hypothesized, positively affect costs. A variety of variables are available here. We use admissions per patient day (ADMISD) and the percentage of patients who are discharged dead (DEATH) as measures of case flow. ADMISPD is computed as the number of yearly admissions divided by total annual patient days; it can be shown that the coefficient of this variable estimates the fixed cost per admission.[1]

---

[1] Let AC = average cost; TC = total cost; x is a vector of explanatory variables, AD = admissions, and TPD = total patient days:

$$AC = a + eb_i x_i + b_n \, AD/TPD$$

$$TC = a \, TPD + eb_i x_i TPD + b_n AD$$

DEATH is calculated as the number of patients who died in the home over the year divided by total annual patient days.[1] Other possible variables not included here include measures of average length of stay and total discharges. As discussed in the footnote, inclusion of the other two variables would lead to overidentification.[2] The heuristic reason is that length of stay is approximately equivalent to the number of admissions divided by patient days, assuming that admissions are approximately equal to discharges and occur at a relatively stable rate.

---

[1] In addition to being a measure of case flow, DEATH may be considered an additional indicator of patient debility or health status. On the other hand, it also might reflect the consequences of lower "quality" care.

[2] Occupancy (OCC) is identically equal to the census (CENSUS) divided by the bed size (BEDS)--i.e.,

(1) OCC = CENSUS/BEDS

If OCC and CENSUS are both included in the cost equation, the BEDS variable cannot be.

Next, in equilibrium the average length of stay (LOS) would be identically equal to the bed size divided by the annual number of admissions (ADM)--i.e.,

(2) LOS = BEDS/ADM

Since BEDS is already determined in (1), either LOS or ADM, but not both, can be included

Furthermore, in equilibrium the annual number of admissions will be equal to the annual number of total discharges (TOTDIS)--i.e.,

(3) ADM = TOTDIS

If LOS has been included, neither ADM nor TOTDIS can be included. If LOS has not been included, either ADM or TOTDIS can be used, but not both.

Finally, the total number of discharges is equal to the number of deaths (DEATHS) plus the number of non-death discharges (OTHDIS)--i.e.,

(4) TOTDIS = DEATHS + OTHDIS

Either of the right-hand variables can be included, but not both.

In summary, we have four equations and eight variables. Thus, only four of the eight variables need to be included. However, as shown above, restrictions emerge from consideration of the individual equations--i.e., it should not be just any four of the eight variables.

Although the above discussion assumes unrealistically that the nursing homes are in dynamic equilibrium, the relationships are nevertheless likely to hold on the average--which is really all that counts anyway in regression.

Input Prices. Costs are hypothesized to be positively related to input prices. Homes located in areas where input prices are higher (e.g., urban areas) also are hypothesized to have higher costs. Ideally, data on the average input prices in the market area (e.g., county) of each facility should enter the cost function since such measures are somewhat exogenous to individual facilities. Unfortunately this type of detailed data is not available. The closest proxy is average nursing wage data constructed by the Department of Labor for only a few SMSAs. In our model input prices are proxied by a group of three variables identified below. First, a location dummy variable was included for nursing homes located in SMSAs (SMSA = 0, if the home is located in an SMSA, 1 if other). Second, data available from the Area Resource File on the average hourly retail age in the facility's county (RETAILING) were included as a substitute for the unavailable nursing wage data.[1] Third, a variable approximately equal to the average hourly wage paid to nurses in the home (NURSEWG) was included in the model.

Regulations. It was argued above that certification, ownership and other direct regulatory variables reflect differences in homes not necessarily related to "quality" or patient condition differences. Because governmental regulations affect the production possibilities of a nursing home, the method of rate determination, staffing requirements, and various payment ceilings, etc., should be associated with cost differentials.

The effects of alternative rate setting approaches cannot be established a priori. On one hand, certain rate setting approaches (in particular, prospective reimbursement mechanisms which pay according to a pre-set schedule) are typically believed to generate incentives to reduce costs. Thus, one might hypothesize that prospective reimbursement mechanisms will be associated with lower operating costs. However such an hypothesis is time-oriented and does not necessarily apply to a cross section of homes across different states. Because rate setting approaches to reimbursement are an historical response by the state to its nursing home expenditures,

---

[1] Preliminary regressions considered two other county level variables describing input prices: manufacturing and construction wages. Use of all three area input price variables resulted in multi-collinearity. Also considered in preliminary regressions were other county-specific variables, including the population density, the number of persons over 65, and the availability of alternative long-term care facilities, be they nursing homes or lower-level facilities (e.g., rest homes). None of the estimated coefficients even approached the magnitude of their estimated standard errors.

reimbursement mechanisms believed to contain costs (e.g., prospective reimbursement) may be found in states with high costs. The model allows for the type of payment mechanism as follows. If the facility is in a state with a "cost related" reimbursement mechanism (e.g., whether it is "actual" or "restrospective" costs), then PM1 is set to 1, and 0 otherwise. If the facility is in a state with a "flat rate" system, then PM2 is set to 1, and 0 otherwise. Thus, the coefficients PM1 and PM2 relate the cost impact of reimbursement systems relative to the omitted category of "prospective" reimbursement. A more complete definition of these variables is contained in Appendix A.

Dummy (1,0) variables are used to indicate whether the facility is in a state with or without other types of regulations. The presence of certificate of need programs is indicated by CON; the presence of occupancy limits is indicated by OCCUPEN;[1] the requirement of a pre-admission patient medical assessment is indicated by MA; the presence of life safety codes is indicated by LSC; the requirement of maintaining a uniform chart of accounts is indicated by UNIF; the presence of limits on administrative salaries and/or total payments is indicated by LIM; and the type of cost basis is indicated by CB, when the rates are set on an individual facility basis as opposed to some form of grouping. Two types of staffing regulations are accounted for: the requirement that a registered nurse or licensed practical nurse be present irrespective of the size of the home is indicated by STAFFA; the number of nurses required per patient day (adjusted for the number of SNF and ICF patients in the facility on the day shift) is indicated by STAFFB.

With regard to certification level, because the Medicare and Medicaid programs impose requirements on SNF and ICF homes which are in accordance with Certification type whether or not appropriate for their patient mix, it is hypothesized that SNF homes (with their greater requirements) have greater average costs than ICF homes. Because mixed SNF-ICF homes must operate, for the most part, according to SNF standards, it is hypothesized that mixed SNF-ICF homes will have costs more comparable to pure SNF than to pure ICF homes. The dummy variables, ICF and ICF-SNF, indicate whether the facility is an intermediate care facility, or a

---

[1] There is no data on the specific level of the occupancy requirement.

combined skilled/intermediate care facility, as opposed to a skilled or extended care facility.[1]

Regional location, reflecting region-specific attitudes to regulation as well as style of care, is indicated by the dummy variables NORTHEAST, SOUTH, and CENTRAL, corresponding to whether the facility is in the Northeast, South, or North Central Census regions, as opposed to the West (the omitted category).

The proportion of patients by payment source indicates more information on the degree to which the facility is subject to regulatory restrictions. The effect of payment mix is not clear. It can be argued that Medicare, relative to Medicaid, imposes cost-generating requirements on homes to care for their more debilitated patients, such that homes with a concentration of Medicare patients relative to Medicaid and private patients have higher average costs. On the other hand, private patients may demand more services, and hence have higher costs than public patients. Since these arguments rely on patient mix and the appropriateness of services to that mix, they suggest that case mix and services are the critical cost determinants given patient mix. Thus, it is hypothesized that payment mix does not affect average costs once account has been made for case mix and services. Patient mix by payment source is controlled for by entering the percentage of private patients (PCTPRI) and the percentage of Medicare patients (PCTCARE) into the model.

Additional Facility Characteristics. In addition to the above types of variables, there remains an *ad hoc* set of variables which reflect managerial decisions about techniques to get on the lowest cost curve. These include: the percent of operating costs which are contracted out to other firms or individuals, PCONTRACT; a dummy (1,0) variable indicating whether the facility is part of a chain of nursing homes, CHAIN; and a variable denoting the magnitude of changes in bed size from the previous year, CHANGE.

---

[1] The certification dummies are undoubtedly picking up some differences in patient mix and service intensity, which have already been accounted for in the cost function. They are present in the model in order to measure cost differences associated with the requirements for different certification types (e.g., standards for the number of differing degrees of paper work associated with each type of certification), regardless of the patient and service mix in the home.

With regard to ownership, some observers feel that non-profit homes provide care of greater amenity levels (e.g., more intensive care) to a different patient mix (e.g., "sicker" patients) . Ownership may also be seen as a variable describing the ability (or desire) of the facility to minimize costs (e.g., non-profit operators may be less aggressive in attempting to achieve efficient production). Some states have implicitly fostered such cost differentials by using ownership as a grouping mechanism to allow different rates for profit and non-profit homes. The (1,0) dummy variable PROFIT indicates that the home is a for-profit facility, as opposed to a non-profit (NONPROFIT) relative to the omitted group, government facility. While these variables are considered here as facility characteristics, taxonomically they also may be considered as output or regulatory variables. The following table, Table T-4.2, summarizes the variables used in this analysis.

T-4.2.3 Estimation Techniques

Consider now the parametric form for estimation of the cost models. There is virtually an unlimited number of ways in which the data might be analyzed, both with respect to the parametric forms to be estimated and the variables to be included. Because of the potential complexity, expense and intractability of reporting a broad range of alternative parametric and variable specifications, we have sought to pursue a single, reasonably general empirical strategy.

First of all, the parametric form estimated should be amenable to manipulation for formula reimbursement. While more sophisticated users would have no difficulty with non-linear alternatives, a formula developed from a non-linear realtionship might be considerably less intuitive for some potential users. Furthermore, in the absence of prior information on the functional form it has now become standard to estimate a linear approximation. Moreover, exploratory investigation of log-linear alternatives did not yield a superior statistical model--i.e., the linear model gave a similar fit.[1] Hence, results reported here are linear ordinary least squares (OLS) estimates.[2]

---

[1] See Section T-4.3 for a discussion of this result.

[2] The basic premise behind the OLS regression analysis is that all firms operate on the same cost curve with the exception of a stochastic disturbance term which reflects purely random cost components (e.g., entrepreneurial vigor). With regard to technical statistical assumptions, it is assumed throughout that the disturbance term has zero mean and constant variance across observations, and also that disturbances across observations are uncorrelated with one another. Moreover, the matrix of independent variables is assumed to be nonsingular and uncorrelated with the error term.

Table T-4.2

VARIABLE DEFINITIONS: NATIONAL ANALYSIS

AOC	Operating costs per patient day, i.e., all annual non-capital costs per patient day that could be avoided only by completely closing the nursing home divided by the number of annual patient days provided
AD	Patient days provided in 365 day units, i.e., average daily census
OC	Percentage of beds occupied by patients
AGE	Average age of patients in the home
FEMALE	Percentage of female residents on night prior to the survey
Diagnosis	
ARTERY	Percentage of patients with diagnosis: hardening of the arteries
MENTAL	Percentage of patients with diagnosis: mental disorders (e.g., mental retardation, psychoses, neuroses, mental illness, emotional problems)
STROKE	Percentage of patients with diagnosis: stroke
ACCIDENT	Percentage of patients with diagnosis: accidents, poisonings, and violence
NUTRIT	Percentage of patients with diagnosis: endocrine, nutritional, and metabolic diseases
OLD	Percentage of patients with diagnosis: senility, old age, and other ill-defined conditions
RESPIRAT	Percentage of patients with diagnosis: respiratory system
DIGEST	Percentage of patients with diagnosis: diseases of the digestive system
NEOPLASM	Percentage of patients with diagnosis: neoplasm
HEART	Percentage of patients with diagnosis: heart attack

Mental Status

DEPRESS	Percentage of patients/depressed
NERVOUS	Percentage of patients/agitated and aggressive
ABUSE	Percentage of patients/abusive and aggressive
SENILE	Percentage of patients confused or senile
SLEEP	Percentage of patients with disturbed sleep

Activities of Daily Living

GROOMPD	Percentage of patients who need help in getting and/or putting away equipment or who need help with bathing but can wash face and hands, brush teeth, comb hair, and shave or apply makeup
GROOMD	Percentage of patients who need help in at leat two of the following activities: washing face and hands, brushing teeth or dentures, combing hair, and shaving or applying makeup
DRESSPD	Percentage of patients who need partial assistance in dressing
DRESSD	Percentage of patients who stay partly or completely undressed
EATPD	Percentage of patients needing major assistance to eat
EATD	Percentage of patients receiving either intravenous feeding or tube feeding
WALKPD	Percentage of patients receiving nurses' assistance in going to the toilet room (do not count on use of cane, walker, or other object of support), using bedpan or commode at night, or cleaning self or arranging clothes after elimination
WALKD	Percentage of patients unable to go to the toilet room for the elimination process
BEDPD	Percentage of patients who need assistance moving in and out of a bed or chair
BEDD	Percentage of patients who are bedfast
CONTINPD	Percentage of patients who are in control of bladder only or bowels only
CONTIND	Percentage of patients who are not in control of bowel and bladder

Output: Services

OT	1 = Occupational therapy offered on premises 0 = Otherwise
RT	1 = Recreational therapy offered on premises 0 = Otherwise
REHAB	1 = Rehabilitation services other than those indicated by PT, OT, RT, SPHT, or COUNSEL are offered on premises 0 = Otherwise
EDUC	1 = Presense of an educational program 0 = Otherwise
PT	1 = Physical therapy offered on premises 0 = Otherwise
SPHT	1 = Speech and hearing therapy offered 0 = Otherwise
COUNSEL	1 = Counseling by social worker, psychologist, or mental health worker offered 0 = Otherwise
NASAL	Percentage of patients in the past seven days receiving nasal feeding
BLOOD	Percentage of patients in the past seven days receiving blood pressure reading
ENEMA	Percentage of patients in the past seven days receiving enema
CATH	Percentage of patients in the past seven days receiving catheterization
BATH	Percentage of patients in the past seven days receiving full bed bath
BOWEL	Percentage of patients in the past seven days receiving bowel and bladder retraining
OXYGEN	Percentage of patients in the past seven days receiving oxygen therapy
BAND	Percentage of patients in the past seven days receiving application of dressing or bandage

PULSE — Percentage of patients in the past seven days receiving temperature-pulse respiration

HYPO — Percentage of patients in the past seven days receiving hypodermic injection

IRRIG — Percentage of patients in the past seven days receiving irrigation

INTRAV — Percentage of patients in the past seven days receiving intravenous injection

SUMSER — Average number of services received per patient per month where the services examined include PT, RT, OT, REHAB, SPHT, and COUNSEL

PONEB — Percentage of rooms with one bed

PTHREEB — Percentage of rooms with three beds

Output: Case Flow

ADMISPD — Number of patients admitted in 1972 divided by annual patient days

DEATH — The number of patient deaths in 1972 divided by the census on the night prior to survey

Input Prices

RETAILWG — 1973 average retail wage in the nursing home's county (taken from Area Resource File)

NURSEWG — Average wage of housekeeping and maintenance personnel (maid, laundryman, maintenance man, etc.) in the home

SMSA — 1 = Not located in SMSA
0 = Located in SMSA

Regulatory Variables

ICF — 1 = Type of certification is ICF only
0 = Otherwise

ICF-SNF — 1 = Type of certification is Medicare & ICF, SNF & ICF, or Medicare, SNF, & ICF
0 = Otherwise

PCTPRI — Percentage of patients whose primary source of payment is own income or family support (private plans, retirement funds, social security, etc.)

PCTCARE — Percentage of patients whose primary payment source is Medicare (Title XVIII)

PM1 — 1 = Type of payment mechanism was cost related (either reasonable, actual, or retrospective)

PM2 — 0 = Otherwise
1 = Type of payment mechanism was flat rate (incl. pts.)
0 = Otherwise

CON — Indicates status of the Certificate of Need Program
1 = Functioning 1972
0 = Not functioning

LSC — Life Safety Codes were in effect in 1972
1 = Yes
0 = No

LIM — Indicates presence of absolute limits on administrative salaries and/or payments
1 = Yes
0 = No

STAFFA — Indicates the requirement that an RN or LPN be present irrespective of the size of the home
1 = Yes
0 = No

STAFFB — Number of nurses required per (SNF/ICF adjusted) patient day for the day shift

MA — Indicates time of patient assessment relative to admission
1 = Pre
0 = Post

UNIF — Indicates the requirement of maintaing a uniform chart of accounts
1 = Yes
0 = No

CB — Indicates type of cost basis (AMS)
1 = Individual facilities
0 = Groups

OCCUPEN — Indicates presence of occupancy penalties
1 = Yes
0 = No

NORTHEAST — 1 = Northeast location
0 = Otherwise

CENTRAL — 1 = Northcentral location
0 = Otherwise

SOUTH — 1 = Southern location
0 = Otherwise

Facility Characteristics

PROFIT — 1 = Individual, partnership, or corporate ownership
0 = Otherwise

NONPROFIT — 1 = Church-related or nonprofit corporate ownership
0 = Otherwise

CHANGE — 1 = Bed size of home changed in past 12 months
0 = Otherwise

CHAIN — If the home is not government owned, this variable indicates whether facility is a member of a group of homes operating under one general authority or general ownership
1 = Yes
0 = No

PCONTRACT — Percentage of operating costs met by contracting with outside group

PAYEXP — Total payroll expenses per patient day

NUREXP — Nursing staff payroll expenses per patient day (includes RNs, LPNs, practical nurses, aides, orderlies, student nurses, and other nursing staff)

FOOD — Outlays for food and other dietary items (include non-contracted services only) per patient day

MAINT — Outlays for purchased maintenance of buildings, grounds, and equipment per patient day

LAUN — Laundry and linen expense (include non-contracted services only) per patient day

COOKWG — Average wage across dietitians, nutritionists, and food service personnel (cook, kitchen help, etc.) in the home

However, the model is not strictly linear with respect to a few "key" variables. Much previous cost modeling--both with respect to nursing homes as well as hospitals--has been concerned with economies and diseconomies of scale and with potential non-linearities with respect to occupancy. Hence, it is desirable that the functional form of the average operating cost function be capable of testing the hypothesis of declining average costs as the output measure (average daily census) increases and as the occupancy rate increases. Such a formulation allows for a test of the existence of economies of scale and utilization. Rather than impose an arbitrary parametric form on the relationship of these variables, the model uses piece-wise estimation, also called "grafting," for this purpose.[1] This procedure, for the most part ignored within the literature, is perfectly general in the sense that it is empirically-based and permits the data to indicate a piecewise linear approximation to an arbitrary, but unknown, functional form.[2] This is accomplished by estimating separate linear segments over different ranges of the explanatory variable as shown in Figure T-4.1. For example, if the explanatory variable in that figure were a scale factor and the dependent variable average operating cost, the results would indicate economies of scale between 0 and X_1, no economies between X_1 and X_2, some diseconomies between X_2 and X_3, and substantial diseconomies above X_3. The technical derivation of this approach is discussed in Technical Paper #3 to this study.

Initially, 10 linear segments were grafted for the measures of average daily census and occupancy. This allowed for determination of a set of join points sufficient to produce a similar approximation with a smaller number of grafted segments. After conducting a series of preliminary analyses, it was determined that six categories were sufficient for both average daily census and occupancy. Investigating alternative join points resulted in the following.

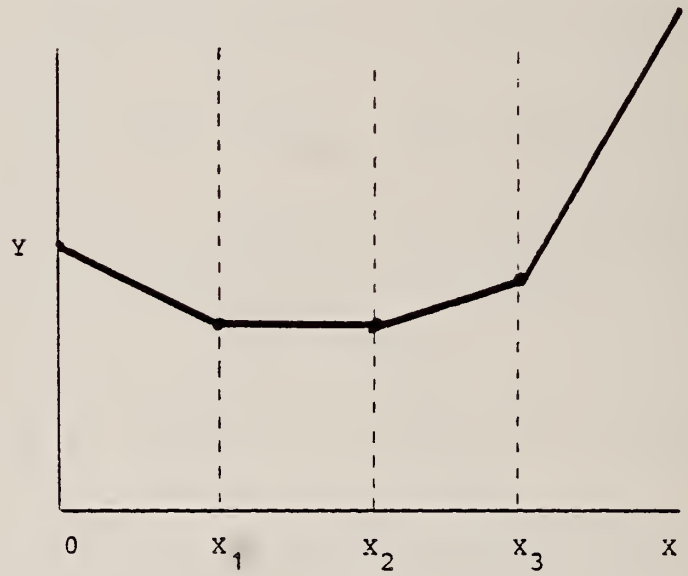
---

[1] For a general discussion of the advantages and application of this approach, see Fuller (1969).

[2] Multicolinearity is also a less serious concern with respect to this alternative. Preliminary results using a quadratic formulation were obfuscated because of severe multi-collinearity in the "occupancy" and "occupancy squared" terms (r = .989) as well as "patient days" and "patient days squared" (r = .909).

Figure T-4.1

ILLUSTRATION OF PIECE-WISE ESTIMATION



For occupancy, separate slope coefficients are obtained for each of the following occupancy ranges: (i) 0-80 percent, (ii) 80.001-85 percent, (iii) 85.001-90 percent, (iv) 90.001-93 percent, (v) 93.001-97 percent and (vi) greater than 97 percent. Such estimation involves the specification of six separate occupancy variables (OC1-OC6) with each defined so that in the appropriate regions they equal either a) zero, or b) the difference between actual occupancy and the upper-bound for the next lower segment, or c) the difference between the upper and lower bound.

For example, if the variable, OC2, is constructed so that it a) equals zero for actual occupancy rates less than 80, b) equals the occupancy rate minus 80 for rates between 80 and 85, and c) equals 85 minus 80 for occupancy rates greater than 85, then the estimated coefficient of OC2 will be the slope of the grafted segment in the range 80 to 85. (See Technical Paper #3 for a precise derivation of the appropriate columns of the data matrix when grafting linear segments.) Thus, if a facility's occupancy rate were 84 percent, then OC1 through OC2 would take the values, 0, (84-80), (90-85.001), (93-90.001), (97-93.001), and (100-97.001).

For the scale factor, we constructed six variables, ADC1 through ADC6 (average daily census (0-20; 20.001-40; 40.001-100; 100.001-150; 150.001-200; 200+) which allowed for estimation of six grafted segments in average daily census as cost varied.

The list of independent variables is lengthy and it should be apparent that the potential for multi-collinearity is great. However, the reader is reminded that multi-collinearity merely leads to inefficiency of some parameter estimates and does not affect the "goodness-of-fit." If our purpose were only to explain cost variations using the available data, even severe multi-collinearity might not be a problem. However, instability of coefficient parameter estimates may be troublesome to users of a prospective reimbursement formula developed from such regression results. Although we have been sensitive to the potential for multi-collinearity in specifying the model--especially with respect to the specifications for average daily census and occupancy--there is no fully satisfactory alternative for resolving the difficulty. This is not a "fishing expedition," for there is a sound conceptual rationale for including each of the variables indicated above.

However, we did investigate the degree of multicolinearity among the variables and eliminated some few variables which were highly colinear and captured elsewhere (e.g., bed size and average daily census; also see footnote x, page xx). Also the piece-wise grafting technique eliminated other problems of multicolinearity. Finally, we not only estimate a complete model but also a "limited" (or partial) model which includes only those variables having an F-statistic[1] of one or greater--i.e., having a coefficient estimate larger than its standard error.--(Section T-4.4). (This corresponds to approximately a 0.30 confidence level.) Coefficient estimates and summary statistics were similar in both the complete and limited models.

T-4.3 The Complete Model

The objectives of the econometric cost function analyses are to identify variables associated with variations in operating costs of nursing homes and to determine the magnitude of the association. The summary statistics of the regrssions reveal that the explanatory power of the model developed above is quite high for cross-section data.[2] The coefficient of determination (R^2) shows that the complete model (shown in Table T-4.3) explains around 58.5 percent of the variation in operating costs. Of the 91 explanatory variables, 25 variables (28 percent) are significant at the 0.10 confidence level or higher.

Before turning to an analysis of the individual explanatory variables, we consider the functional form of the equation. For the most part, previous research has concentrated on additive, linear forms for nursing home cost functions. While non-linear logarithic forms may have considerably less

---

[1] Note that an estimated coefficient with an F-statistic equal to 1 has the same statistical significance as if it had a t statistic equal to 1.

[2] For comparison purposes, the AMS (1976) analysis of private patient charges explains around 56-60 percent of the variance.

Table T-4.3: DETERMINANTS OF AVERAGE OPERATING COSTS: THE COMPLETE COST MODEL; NATIONAL NCHS DATA

Explanatory Variable	Estimated Coefficient	Standard Error
Output: Scale		
CONSTANT	24.2978	
AD1: 0-20 Beds	-0.9678xxx	0.2143
AD2: 20-40 Beds	-0.0692	0.0442
AD3: 40-100 Beds	-0.0355xxx	0.0098
AD4: 100-150 Beds	0.0173	0.0142
AD5: 150-200 Beds	-0.0242	0.0179
AD6: 200+ Beds	0.0118x	0.0065
OC1	-0.0132	0.03726
OC2	-0.1181	0.2054
OC3	0.0883	0.2206
OC4	-0.2209	0.2595
OC5	0.0219	0.1508
OC6	-0.0300	0.1675
Output: Patient Mix		
AGE	-0.0137	0.0358
FEMALE	-0.0093	0.0124
Diagnosis		
ARTERY	-0.0261xx	0.0113
MENTAL	-0.0166	0.0142
STROKE	0.0227	0.0148
ACCIDENT	0.0267	0.0210
NUTRIT	0.0260	0.0240
OLD	0.0017	0.0130
RESPIRAT	0.0179	0.0311
DIGEST	0.0347	0.0332
NEOPLASM	0.0067	0.0313
HEART	0.0061	0.0180
Mental Status		
DEPRES	0.0150xx	0.0076
NERVOUS	0.0010	0.0083
ABUSE	-0.0129	0.0114
SENILE	0.0045	0.0079
SLEEP	-0.0109	0.0088
Activities of Daily Living		
GROOMPD	0.0086	0.0096
GROOMD	-0.0131	0.0126
DRESSPD	0.0061	0.0125
DRESSD	-0.0050	0.0153
EATPD	0.0195	0.0125
EATD	0.0683	0.0509
WALKPD	0.0105	0.0115
WALKD	0.0104	0.0142
BEDPD	-0.0027	0.0091
BEDD	0.0228	0.0214
CONTINPD	-0.0404xx	0.0181
CONTIND	-0.0096	0.0115

Table T-4.3: DETERMINANTS OF AVERAGE OPERATING COSTS: THE COMPLETE COST MODEL; (Cont'd.) NATIONAL NCHS DATA

Explanatory Variable	Estimated Coefficient	Standard Error
Output: Services		
OT	0.8629xx	0.3959
RT	0.4081	0.3595
REHAB	-0.3649	0.5855
EDUC	0.4643	0.3224
PT	0.9129xx	0.4531
SPHT	0.0511	0.5040
COUNSEL	0.2547	0.3787
NASAL	-0.0348	0.0549
BLOOD	0.0077	0.0053
ENEMA	0.0168x	0.0101
CATH	0.0188	0.0250
BATH	-0.0040	0.0057
BOWEL	-0.0030	0.0085
OXYGEN	-0.0470	0.0698
BAND	0.0142	0.0161
PULSE	0.0059	0.0051
HYPO	0.0039	0.0073
IRRIG	0.0024	0.0198
INTRAV	0.0040	0.0004
SUMSER	0.0756xxx	0.0267
PONEB	0.0133	0.0106
PTHREEB	-0.0017	0.0180
Output: Case Flow		
ADMISPD	419.3255xxx	43.7281
DEATH	0.4914	0.8400
Input Prices		
RETAILWG	0.9781xxx	0.3685
NURSEWG	2.7800xxx	0.1962
SMSA	-0.4922	0.3647
Regulatory Variables		
ICF	-2.1250xxx	0.4822
ICF-SNF	-0.4556	0.3939
PCTPRI	0.0024	0.0069
PTCARE	0.0324x	0.0187
PM1	-1.4361xxx	0.5313
PM2	-1.2972xx	0.5698
CON	0.0963	0.5983
LSC	0.7719x	0.3988
LIM	-0.1195	0.6513

Table T-4.3: DETERMINANTS OF AVERAGE OPERATING COSTS: THE COMPLETE COST MODEL; (Cont'd.) NATIONAL NCHS DATA

Explanatory Variable	Estimated Coefficient	Standard Error
Regulatory Variables (continued)		
STAFFA	1.4634xx	0.6987
STAFFB	0.3790	1.8280
MA	0.3258	0.6070
UNIF	0.7543	0.5315
CB	-0.0934x	0.4413
OCCUPEN	0.9779x	0.5464
NORTHEAST	3.6221xxx	0.9301
CENTRAL	1.7824xx	0.7814
SOUTH	2.6091xxx	0.7728
Facility Characteristics		
PROFIT	-2.5167xxx	0.6566
NONPROFIT	-0.7752	0.6941
CHANGE	0.6598	1.1558
CHAIN	-0.4917	0.3108
PCONTRACT	0.0359	0.0363

R^2	0.5847
Adjusted R^2	0.5842
F-ratio	16.0146
N	1127
Standard Error	4.79

x significant at .10 level
xx significant at .05 level
xxx significant at .01 level

intuitive appeal for non-econometrician rate settors,[1] if such non-linear forms give considerably better results than more "understandable" linear regressions, they may be justifiable. It can be shown that a multiplicative average cost function is appropriate for a Cobb-Doublas production function and should be considered. Results for a representative multiplicative function are presented in Appendix T-4.A, Table T-4.6. The dependent variable in that equation is the natural logarithm of average operating expense. In most cases, the same individual coefficients are significant for both linear and nonlinear forms. The next comparison is to consider the relative explanatory power of the two types of equations. To compare the explanatory powers, it is necessary to compute and square the derivation of the predicted values of operating expenses. These squared residuals are then compared to the sum of squared residuals in the linear equation for the same date. This computation reveals an estimated R^2 for the semilog form of 0.4517, compared to the R^2 for the linear form of 0.4625. Hence, the linear and semilog forms perform about the same in terms of explained variation in operating costs; indeed the linear form performed slightly better.

In light of the absence of any gain in model accuracy in using the semilog form and the increased complexity in implementing the semilog form, we consider only the linear form in the remainder of this analysis.

---

[1] The regression coefficients for the semilog form are not as readily interpretable as the coefficients in the linear form. As shown below they can be interpreted in terms of elasticities. If σ_{xi} is the elasticity of average costs with respect to X_i, then σ_{xi} measures the proportional rate of increase of average costs for proportional increases in X_i from the level in question, i.e., the percentage change in average costs resulting from a one percent change in X_i from the level in question. For the semilog specification,

$$\sigma_{xi} = \frac{\partial \text{ OPEXP}}{\partial x_i} \frac{x_i}{\text{OPEXP}} = \frac{\beta_i \exp [\Sigma \beta_i X_i + u] X_i}{\exp [\Sigma \beta_i X_i + u]}$$

$$= \beta_i X_i$$

In measuring the effects of various explanatory variables on operating costs, the use of elasticities is cumbersome since it entails measuring a different impact at various levels of the explanatory variables in question, and the measured percentage changes in costs should then be converted to absolute changes.

T-4.3.1 Output Measures

T-4.3.1.1 Scale Effects

However, as discussed in Section T-4.2.3, the "grafting" approach allows the model to be non-linear with respect to the key output variables of average daily census (ADi) and occupancy (OCi). The grafted average daily census variables have statistically significant negative coefficients[1] for the ranges of 0-20 (AD1) and 40-100 (AD3); thus it appears that there are returns to higher capacity over these ranges. Although the coefficients of AD4 (100-150) and AD5 (150-200) are statistically insignificant, the positive coefficient of AD6 (200 or greater) is statistically significant though its absolute magnitude is slight. For example an increase in AD from 200 to 250 patients will lead to a cost increase of $0.50--or 1 cent for every additional patient day. The graph in Figure T-4.2 displays these ADC coefficients and shows a minimum cost home to have an average daily census in excess of 100 patients, since beyond AD$\geq$40, the cost curve is virtually flat.

With respect to occupancy, the other non-linear variable, it appears that there are no returns to higher levels of utilization. All OC variables are insignificant individually, or as a group.[2] Figure T-4.3 displays these coefficients. In combination with the AD results, this suggests that the optimal size of a nursing home is in excess of 100 beds.[3]

---

[1] Unless otherwise noted the use of statistical "significance" in the text of the discussion refers to the 0.10 levels; the tables reporting the regression results denote levels of significance at 0.01, 0.05, and 0.10. Hence, 0<AD<20 and 40<AD<100 are significant at the 0.01 level, while the coefficient of AD>200 is significant at the 0.10 level.

[2] An F-test was performed on the hypothesis that the occupancy rate coefficients, when taken together, were each equal to zero. The test statistic,

$$\left[\frac{23866.88-23784.19}{23784.19}\right] \times \left[\frac{1127-82}{6}\right]$$

$$= .559727$$

is distributed as an F(6,1035); it is insignificant at even the 25 percent level (F(.25,6,1035) 1.32). Thus we fail to reject the hypothesis that (after controlling for output, input prices, regulation, and other facility characteristics), there are neither economies nor diseconomies of utilization. In other words, the OCi coefficients have no explanatory power. However, they are maintained in the regression for completeness.

[3] It should be noted that preliminary analyses of alternative specifications of the average daily census and occupancy variables in terms of their more common quadratic formulation affected neither the magnitude nor the significance of the other coefficients. However, the coefficients of AD, AD-squared, OC, and OC-squared were plagued by the multi-collinearity (their correlation coefficients approached 0.95).

Figure T-4.2

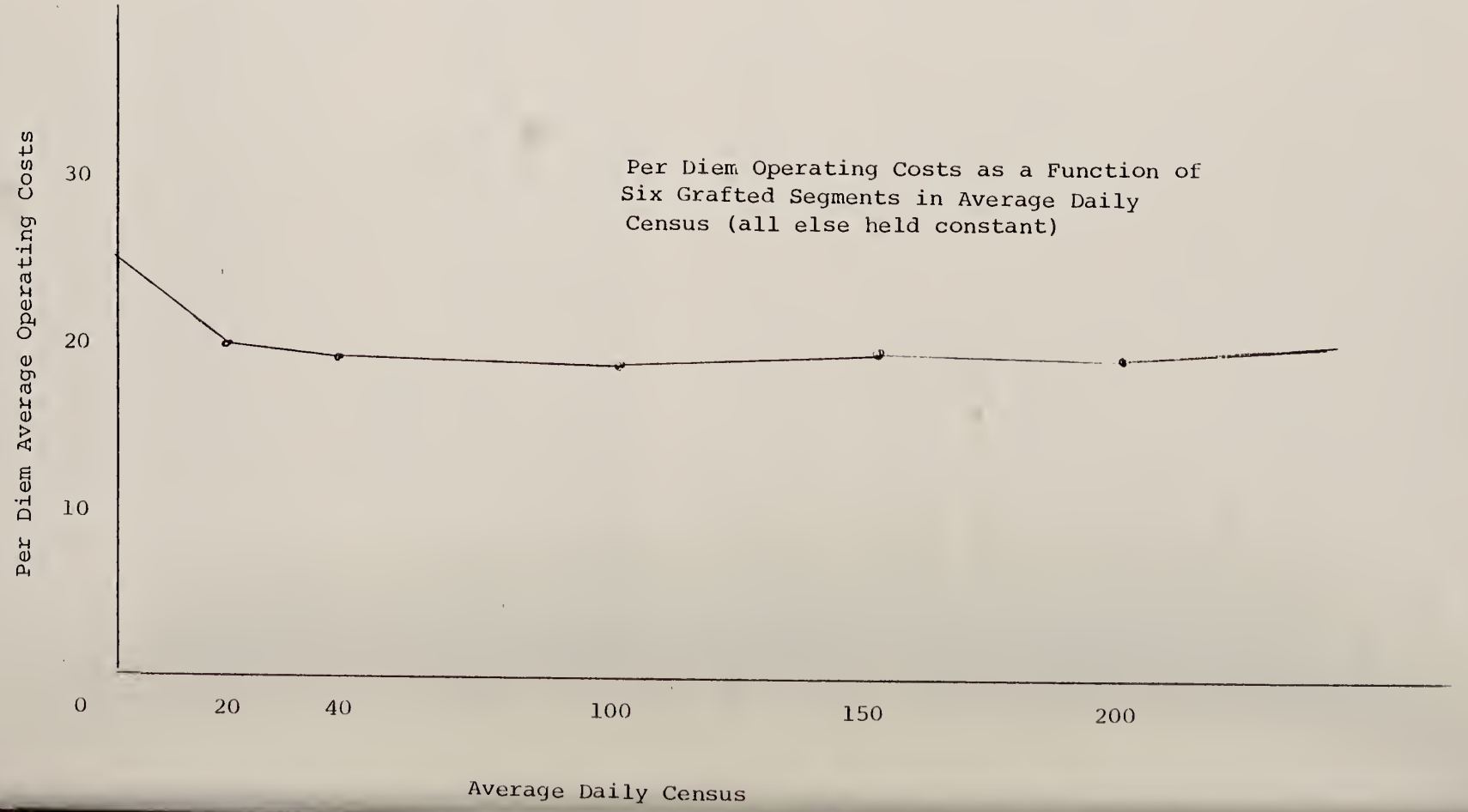
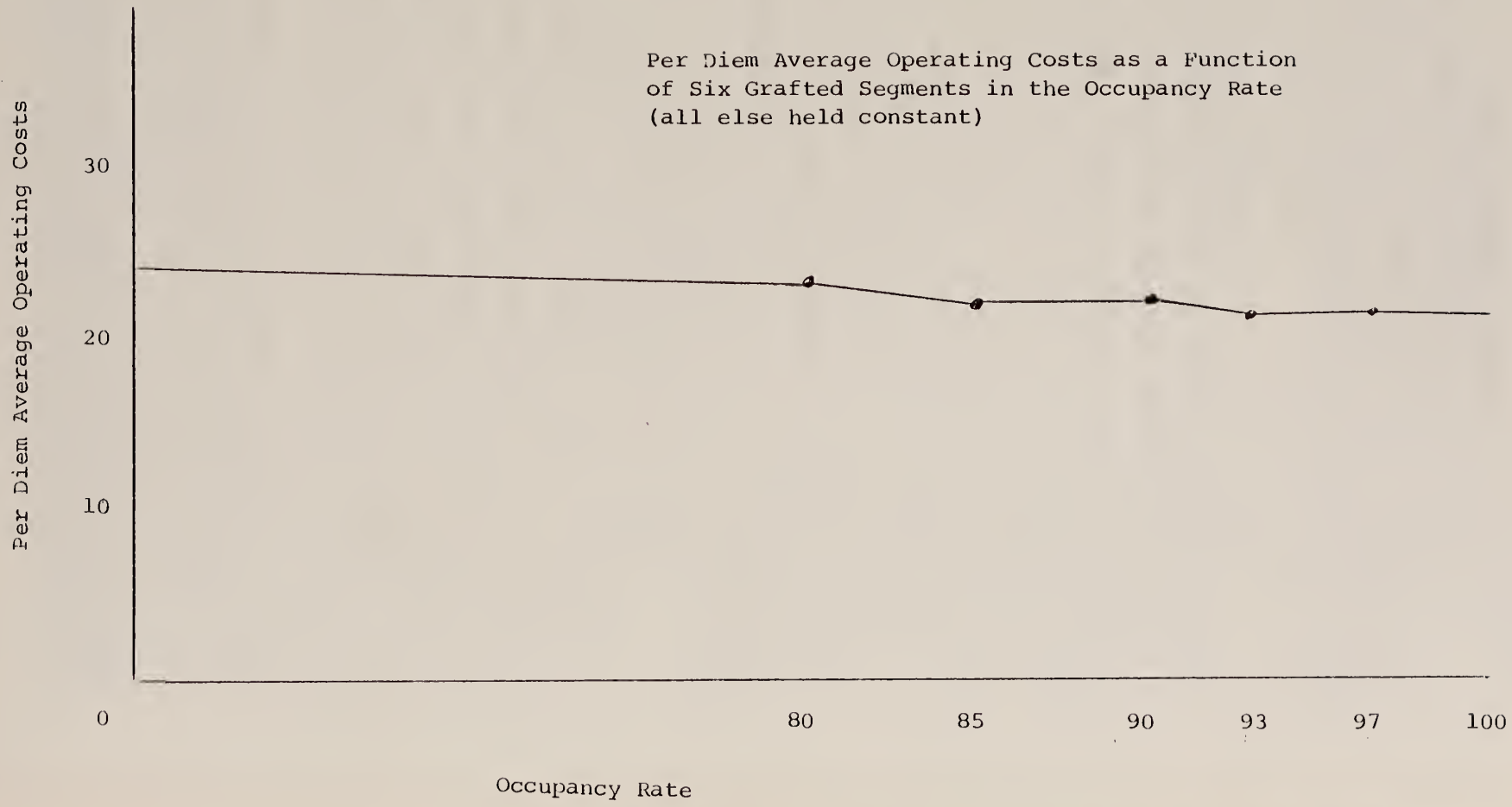

Per Diem Operating Costs as a Function of Six Grafted Segments in Average Daily Census (all else held constant)
Per Diem Average Operating Costs
30
20
10
0
20
40
100
150
200
Average Daily Census

Figure T-4.3


Per Diem Average Operating Costs as a Function of Six Grafted Segments in the Occupancy Rate (all else held constant)
Per Diem Average Operating Costs
30
20
10
0
80
85
90
93
97
100
Occupancy Rate

The remainder of this section discusses the other explanatory variables.

T-4.3.1.2 Patient Characteristics

A wide range of patient mix variables are available from the NCHS data: age, sex, activities of daily living, measures of debility, measures of diagnosis and mental condition at the time of the survey. In general, these variables contribute little to the model. Both the age and sex variables were insignificant.[1]

Among the patient diagnosis variables, only the percentage of patients with hardening of the arteries (ARTERY) was statistically significant, although the percentage of patients with mental disorders (MENTAL), cardiovascular disease (STROKE), accidents (ACCIDENT), or endocrine or metabolic disease (ENDOCRIN) were insignificant. Note though that all diagnostic coefficients are in the correct (i.e., positive) direction. With regard to mental status, only the positive coefficients of the percentages of patients who were depressed (DEPRES) achieved statistical significance. The 1 cent to 2.6 cent effect on cost per percentage of patients in these several categories would be of economic significance only if a large percentage of the patients in the home had the diagnoses. For example, if 24 percent of the patients in a home had hardening of the arteries (the mean percentage in the sample) per diem operating costs would rise around $0.50(0.02x24).

Patient Characteristics: Debility

Among the activities of daily living (ADL) measures of debility, the only variable with a statistically significant coefficient was the percentage of patients who were partially incontinent (CONTINPD). Its negative sign was surprising; perhaps patients who are in complete control of their bladder and bowels require more nursing care in the form of supervision and planning of recreational events. While the coefficient was small (4 cents per percentage of patients partially incontinent), if a large portion of the home were incontinent, the overall effect would be large. However, on

---

[1] The insignificance of these patient descriptors may be a result of the sampling procedure used in the data. Some number of patients less than 10 were sampled in each home. Thus the difference between the mean age of those patients sampled and actual mean age is probably a small random quantity, and would increase the standard error of the estimated coefficient, making it insignificant. Because the age information is based on a sample of patients less than 10 in quantity, it was not feasible to develop age distribution variables (as was done in the state analyses).

average only 6 percent of the patients in homes in the sample were partially incontinent; on average, then, the per diem cost impact is $0.24. In any event, the regression results suggest that on grooming, dressing, eating, walking, and being bedfast, the costs of caring for different types of patient types do not significantly differ.[1] Of course, these results are likely to reflect multi-collinearity and the limited model reported in Section T-4.4 is less likely to suffer this problem.

Patient Characteristics: Summary

Many persons, especially nursing home operators, believe that the mix of patient conditions is a significant determination of nursing home operating costs. Such beliefs are neither supported or rejected by these results. The previous analysis suggested that few individual patient descriptors were statistically significant. Moreover, their economic magnitude would be large only if a large percentage of the patients in the home were described by these variables. To assess the maximum average overall impact, consider all patient characteristics whose estimated coefficients have F-statistics greater than 1. Note that this implies a very low probability (0.3) that the coefficient is statistically different from zero; we choose this low threshold to include a maximum number of patient characteristics. If we then multiply all of those patient characteristics coefficients by the mean percentage of patients having those characteristics and sum over characteristics, the result is a maximum estimate of the average cost impact of patient characteristics. The result of this exercise is that the maximum average effect of patient characteristics is about $1.05 or less than 8 percent of the mean average operating costs of the homes in the data ($13.53). If we consider only those variables with coefficients statistically significant at the 0.10 level (F>2.71), the resulting cost impact is far lower, $0.21. Thus in the context of the other variables which affect per diem cost, discussed below, patient characteristics have a relatively low impact. However, this may be due to the downward errors-in-variables bias. Note that this maximum average effect may be affected by multi-collinearity. When the same calculation is made for a "limited" model with less multi-collinearity (Section T-4.4), a comparable result is found.

---

[1] The hypothesis that when taken together, the coefficients of each of these variables is zero was tested. The appropriate F-statistic is given by 1.584 which is insignificant at the 0.25 level (F(0.25,11,1036)=1.245).

T-4.3.1.3 Services

Three types of service variables were available to proxy quality of care. The first group is a series of seven dummy variables indicating the type of services offered by a nursing home. The second group is a series of 12 variables describing the intensity of services provided. The third group includes measures of the percentage of patients in different size rooms, as quality measures of a less clinical nature.

Two of the service offered variables were statistically significant. Offering occupational therapy (OT) raised average operating costs by $0.86; offering physical therapy (PT) raised average operating costs by $0.91. The remainder of the service offered variables were statistically insignificant. Except for the rehabilitation variable, all signs, as expected, were positive, indicating that homes offering these services are more costly.

The service intensity variables measure the percentage of patients in the week prior to the survey receiving one of 12 services. Of these service intensity variables, only ENEMA patients achieved statistical significance, though of a low economic magnitude ($0.02). Here too, if a large percentage of patients required an enema, the cost impact could be large. However, since on average 14 percent of the patients in a home required an enema, the average cost impact is around $0.28.

An alternative service intensity variable (SUMSER) was constructed; SUMSER accounts for service intensity by entering the average number of services received per patient during the month prior to the survey.[1] The positive and significant coefficient of SUMSER indicates that on average, it costs about $0.08 to provide one or more doses of any of these services to a patient. Since, on average, patients received 4.1 services per month, the cost per patient day of providing these services is around $0.33.

These results imply that while there are costs to facilities of providing a more intensive level of services, the magnitude of the cost of altering the intensity of a service is small relative to the cost of simply offering the service on an as-needed basis. The inference is that nursing home costs vary in accordance with the presence or absence of offering a service, but not the intensity of its use. This result suggests that nursing homes do not alter their staffing mix or other variables in their control in

---

[1] The set of services examined differs from the 12 accounted for earlier. Considered here are the intensity of physical, recreational, occupational, speech or hearing therapy, as well as professional counselling.

response to patient condition, although they may staff up to what they perceive to be the "steady state" of their patient mix. Costs appear to accrue to the organization of the home and the staff; facilities and staff are there and must be paid, whether or not they provide services.[1]

Finally, the percentages of patients in either one bed (PONEB) or three or more bed (PTHREEB) rooms was included as quality measures. Both coefficients were statistically insignificant.

T-4.3.1.4 Case Flow

The final measures of facility output were admissions per patient day (ADMISPD) and the percentage of patients who died in the home (DEATH). The costs per admission (the coefficient for ADMISPD) was estimated at $419.00 and was significant at the 0.01 level.

T-4.3.2 Input Prices

The average retail wage (RETAILWG) in the county in which the facility was located is used as an index of regional wage levels. The resulting coefficient is positive and highly significant (at the 0.01 level); the magnitude of the coefficient ($0.98) suggests that nursing home costs closely

---

[1] Note, however, that the average cost impact of services offered ($1.08) is about the same as the average cost impact of service intensity ($1.07). This calculation is based on a composite "average" home that provides services of the type and intensity described by the mean of those services offered and service intensity variables whose estimated coefficient has an F-statistic greater than one. Multiplying the estimated coefficients by the mean percentage of homes and patients offering and using the services, the result is that on average services offered increase per diem costs by $1.08, while the average impact of service intensity is $1.07. If we consider only those variables with a coefficient statistically significant at the .10 level (F>2.71), the resulting average cost impact of services offered is $0.86 and the average cost impact of service intensity is $0.47. The "limited" model presented in Section T-4.4 has corresponding results for F>1.0. The average cost impact is $1.45 and $1.50 while for F>2.71, the average cost impact is $1.16 and $1.04 respectively for services offered and service intensity. The meaning of such a calculation is not clear, for "composite" homes do not exist. No doubt, homes offer services in clumps, such that some homes offer many services, while others offer no services. Nor is it clear that these "average" costs are relevant for a reimbursement system designed to induce "marginal" responses from operators. Using such "average" costs corresponds to average cost pricing and induces overproduction of services which are inappropriate to patient mix. While such numbers may be of interest to those bureaucrats anxious to increase the size of the industry they are regulating, they are clearly inappropriate to a "marginal" cost pricing system of the formula prospective reimbursement variety being considered here.

follow the local wage level. The NURSEWG variable measuring the average hourly LPN nurse wage within each facility also was highly significant (indeed, it was the most significant of all the explanatory variables) and positive ($2.79). Urban/rural location (SMSA) also was included as a control for systematic price differences in more urban areas. While the estimated coefficient was not statistically significant, it was larger than its standard error; its negative sign suggests that urban areas may have higher costs than rural areas that are not captured by the wage variables.

T-4.3.3 Regulatory Variables

In general, the regulatory variables add greatly to the explanatory power of the model. Beginning with the certification level variables, intermediate care facilities (ICF) cost approximately $2.12 less per patient day than the omitted skilled nursing facility category; the coefficient is highly significant (at the 0.01 level). Mixed ICF-SNF homes cost approximately 50 cents less per patient day than pure SNFs; however, while the coefficient is larger than its standard error, it is not statistically significant and suggests that mixed ICF-SNF homes follow closely the input structure of pure SNF facilities in order to meet certification standards.

With respect to payment source, it appears that otherwise similar private and Medicaid patients cost about the same (PCTPRI is statistically and economically insignificant), but that Medicare patients cost somewhat more. The estimate of the PCTCARE coefficient implies that the per diem cost in a facility which has only Medicare patients would be about $3.25 higher than a facility with all Medicaid or private patients. This variable, too, may reflect output-quality variations as well as regulatory effects.

Turning to the impacts of specific regulations, it appears that payment mechanism is a significant cost determinant. Two variables indicating whether the reimbursement system was a retrospective cost related (PM1) or a flat rate system (PM2) were entered into the equation. Their coefficients reveal that, ceteris paribus, homes under prospective systems have

higher costs than homes under cost-related systems; cost-related systems differ significantly from prospective systems by about $1.43 per patient day, while flat rate systems differ significantly from prospective systems by about $1.30 per patient day. This result must be interpreted cautiously. On one hand, prospective reimbursement may lead to increased costs, for example, to pay providers a risk premium to cover the possible costs of unforeseen events. Whatever the reason for higher costs under prospective reimbursement, such higher costs apparently do occur, as found by Sloan and Steinwalt (1978)--in the most definitive study yet of the effect of prospective reimbursement. Cromwell (1977) too found no evidence of cost containment under prospective reimbursement. However, while it may be that higher costs are a by-product of prospective systems, another explanation may be that those states with excessively high nursing home costs adopted a prospective system in an effort to restrain cost increases. The higher cost of homes in states with prospective systems would then reflect the generally higher cost structure of these states, and not that prospective reimbursement leads to higher costs. The correct test of the effect of prospective reimbursement would be to observe homes in both prospective and retrospective rate states over time. Note that as discussed, there were several possible sources of data on these reimbursement variables. Although there were minor differences in the assignment of states to a particular reimbursement system, use of each of these data sources in comparable analyses reached in the same result: homes in states with prospective systems cost more than homes in states with cost-based systems. Despite difficulties in depicting the rate setting mechanism imposed in 1972, the robustness of the results under any set of definitions allows confidence in the differential result of payment systems.

The coefficients of the dummy variables indicating the presence of a certificate of need program (CON), preadmission patient assessments (MA), and uniform charts of accounts (UNF) are positive but not significant. That the coefficient of the life safety codes in effect variable (LSC) is positive and significant no doubt reflects the costs of implementing LSC requirements and, in this regression, increases operating costs per patient day by about $0.77. Reimbursement systems which put limits on total payments to nursing homes or put limits on administration salaries may be successful in reducing costs as indicated by the negative coefficient of LIM; however, the coefficient was not statistically significant.

Homes which had an "absolute" staff requirement (i.e., were required to staff an RN or LPN irrespective of the size of the home--STAFFA) had significantly higher costs by $1.46. However a "relative" staff requirement (i.e., requiring a set number of nurses per patient day--STAFFB) did not have a statistically significant effect. As in the case of the service variables, results here appear to indicate the differential cost impact of absolute, as opposed to relative, commitments of resources to meeting requirements. Finally, occupancy penalties (OCCUPEN) significantly and positively affect costs. This result complements the earlier finding that there are no returns to higher levels of utilization. Hence, states which penalize homes for low occupancy levels are apparently penalizing homes for what cannot be shown to be inefficient behavior.

Finally, controlling for regional variation there are highly significant differences in nursing home costs across the nation. Homes in the Northeast region cost about $3.62 more than homes in the omitted West region; homes in the South cost about $1.78 more; and homes in the Central region cost about $2.61 more. These regional differences may reflect regulatory differences, or perhaps differences in general patterns of care, or otherwise unmeasured price differences. However, inasmuch as patient, service, and price differences have already been controlled, we group these regional variations under the rubric of regional regulatory differences, perhaps reflecting local attitudes toward long-term care.

T-4.3.4 Facility Characteristics Describing Production Differences

The final set of variables in the cost function include additional facility characteristics which may affect facility production possibilities and hence cost.

For-profit homes (PROFIT) have significantly lower per diem costs, by $2.52, than do the omitted government homes. While the coefficient on the non-profit (NONPROFIT) variable is statistically insignificant, its coefficient is greater than its standard error; its negative sign together with its relative large magnitude ($0.78) suggests that non-profit homes may have lower costs than government facilities, but certainly cost more than for-profit homes. In any event, for-profit homes cost the least (by a range of $1.74-2.52 per patient day); while their cost differentials may reflect unmeasured quality output differences, the wide range of service and patient characteristics included in the model is the most complete test possible of

this situation, with available data. These results are confirmed in the state-specific analyses. In the absence of additional information, one must conclude that patients in for-profit homes incur lower costs than otherwise comparable patients in non-profit and government facilities.

Another facility related variable is chain ownership (CHAIN). The coefficient of this variable was significant at the 0.15 level but not the 0.10 level. This is evidence that facilities operating under chain ownership cost about $0.49 less per diem than nonchain homes. Variables accounting for changes in bed size over the previous year (CHANGE) and the percent of costs provided under contract (PCONTRACT) were not significant.

T-4.3.5 Alternative Specifications: Grouping

This section investigates the possibility of separate cost curves by certification level to evaluate whether homes providing different levels of care (SNFs, ICFs, and combined SNF-ICFs) have different cost structures. The differences in care provided to the patients in homes of different certification levels may extend to many or all the variables affecting costs, e.g., facility characteristics, services offered to patients, etc. If the individual skill-level cost curves are, in fact, different, then a cost curve based on pooling SNF, ICF and combined SNF-ICF homes into one sample (and allowing only intercept differences in the form of dummy variables for certification level)--as in the previously reported results--would be are inappropriate. Rather, the homes would need to be grouped into separate samples by certification level with separate cost curves estimated for each.

Three subsample regressions were estimated. The explanatory variables were those used in the complete model with the exception that the certification dummy variables were omitted. An F-test was used to assess differences in the cost functions of homes in each of the certification classes. The null hypothesis is that the overall structure of operating costs is the same for homes of each certification class.[1]

---

[1] Specifically, we compare the residual sum of squares from the two model specifications:.

(1) $ADC = \beta_o + Z\beta + U$ (restricted)

(2) $ADC = \beta_o + Z\beta + \sum_{i=1}^{2} \lambda_i X_i A_i = U$ (unrestricted)

This statistic is insignificant at the 0.05 percent level, $F(0.05;863,264)=1.19$. Hence, the hypothesis that the cost functions are the same for different certification types cannot be rejected. Thus we conclude that one cost equation, including all three certification types, is an adequate econometric specification.

The regression results for the three certification subsamples are reported in Appendix T-4.B, Table T-4.B-1. In addition to the test results, we note that the coefficients of the separate estimates do not differ by economically meaningful amounts.

T-4.3.6 Alternative Measures of Debility

It is possible that alternative methods of measuring the patient debility mix of the home yield other results. While the remainder of this section digresses from the main discussion of results, it is important in light of recent interest in ADL measurement.

Several alternatives were investigated. The first alternative followed the Winn, (1975) procedure of constructing a variable to represent the number of nursing hours which should be allocated to the debility needs for the average patient in the home.[1] The coefficient of this variable

---

(continued from prior page)

where Z is the (n x k) matrix of independent variables; X is the (n x (k-2)) matrix of all independent variables except the two certification dummies; β_o (K x 1) and λ_1 and λ_2, each ((K-2)x1) are vectors of parameters; U^o is an (n x 1) vector of disturbances; and A_i is defined so that it equals 1 for sub-sample i and zero otherwise. Notice that specification (2) is equivalent to stratifying the estimation by certification type.

The F test is applied to the change in the sum of the squared residuals (SSE) incurred by the restricted specification (1), i.e.

$$\frac{SSE(R)-SSE(U)}{SSE(U)} \cdot \frac{n-r}{t} \sim F\left[\begin{matrix} t \\ n-r \end{matrix}\right]$$

where SSE(R) = the sum of squared errors estimated under the restriction

SSE(U) = the sum of squared errors estimated without the restriction

r = the number of parameters in the unrestricted equation

t = the number of restrictions

n = the number of observations

Testing for the same cost function for SNF, SNF-ICF, and ICF homes, the calculated F-statistic was 0.8267.

[1]The hourly weights used in the construction of this index are presented in Appendix T-4.A as Table T-4.A.2.

(ADL1) was insignificant and of such small magnitude to be meaningless; other coefficients in the regression remained as before. Table T-4.3 presents the coefficients of this and the other alternative ADL measures.

Table T-4.3

RESULTS OF USING ALTERNATIVE MEASURES OF DEBILITY

	Estimated Coefficient	Standard Error
ADL1	0.0041	0.0026
ADL2	0.0075	0.0086
ADL3	-0.3000	0.4162
ADL4	0.3258	0.4585
ADL5	0.0313	0.6942

The second index considered (ADL2) is conceptually similar but also accounts for nursing hours needed to care for patients with diseases of the nervous system and mental status, while omitting several of the ADL measures.[1] Again the coefficient here was insignificant and other coefficients remained as before. Next we adapt a one-dimensional index employed by Carroll (1962), Wylie (1964), Dinnerstein, et al., (1965), Linn (1967), and Schoening and Iverson (1968). This index is the natural extension of a one-dimensional index used to rank patients in terms of debility, averaged over patients. This patient level index, called a sum-of-scores index (ADL3), is ordinal, so that the ranges of the interval between successive scores are not of equal value in terms of some absolute concept of debility. The patient level index is computed as the arithmetic sum of a patient's scores across the six ADLs examined where, for each ADL, a score of zero corresponds to complete independence, a score of one corresponds to partial dependence, and a score of two corresponds to complete dependence. The mean index score in the home is taken as the measure of the overall debility of the home. As before, the coefficient was not significant and the other coefficients remained the same.

A different but conceptually similar index was constructed by weighing the partially dependent categories with a score of two and the dependent categories with a score of one. If, as suggested by Skinner and

---

[1] See Appendix T-4.B, Table T-4.B-3.

Yett (1973), it is the partially disabled patients who are the most costly, one would expect this alternative sum of scores index (ADL4) to do somewhat better than the first sum of scores index since the higher scores correspond to the more partially disabled patients. This index, too, yielded insignificant results.

The insignificance of both the sum of scores indices is not surprising given that they do not explicitly account for differences in the amount of care (hence costs) associated with different performance levels of the various ADLs, as do the Winn-type indices. By using such indexes, ADLs have been implicitly treated equal in the degree of care associated with them, although ADL4 made some steps in this direction. For rate setting purposes, this is clearly a deficiency if the extent to which the range and the types of disabilities present are related to the costs of caring for the patient and are to be considered.

A final index which was examined is the natural extension of the Guttman index which has been used extensively at the patient level as in Skinner and Yett (1973). At the patient level, the Guttman debility index is identical to the sum of scores index in actual construction if a simple dichotomous categorization scheme, either independent (corresponding to 0) or dependent (corresponding to 1), is used to classify patient performance on each ADL. It is different from the sum of scores index only in terms of the way scores are interpreted. Use of the Guttman debility index is based on the assumption that the probability is very high that as a patient changes from being completely independent to completely dependent, (s)he will do so by becoming dependent in the various ADLs in one pattern only. Put another way, suppose that a total of n ADLs are being examined. Then this procedure implies that out of the n factorial possible patterns of successively incurring disabilities, the probability of a patient possessing the same types of disabilities is very high. For example, in a set of 5 ADLs, we would expect that all individuals who exhibit dependence in 4 ADLs do so with respect to the same 4; that those who exhibit dependency in 3 ADLs do so with respect to the same 3, and so on.

The problems with the Guttman index are first, that it is unrealistic to presume knowledge about which debilities have most likely generated particular scores. Second, the amounts of care associated with different

disabilities have not been accounted for. Third, and most important, it is unreasonable to presume that patients with the same number of disabilities will have the same type of disabilities.

In any event the empirical results speak for themselves; once again this alternative measure of patient debility (ADL5) yields nonsignificant results and does not alter the other coefficients.

Each of the three case-mix measures which were constructed as the averages of patient level indices (ADL3, ADL4, and ADL5) have intuitive appeal but provide only an ordinal measure of debility, i.e., the range of the intervals between successive scores are not of equal value in terms of debility. Therefore, one cannot measure the additional care required (hence costs) associated with changes in debility. In each case, interpreting the values of estimated coefficients is unclear. In addition with ADL1 or ADL2 there is no way to detect which types of disabilities are present. It is important to account for disability types since types may vary in the required nursing care (hence costs) associated with them. Two homes with equal index values may have quite different costs depending on which patient disabilities generate their score. With the use of the 12 variable index (used in the "complete" model discussed prior to consideration of the alternatives) which measures the percentages of patients in the different debility categories across the six ADLs, this problem is avoided; the case-mix variables are cardinal. This property facilitates a clear interpretation of estimation results. Thus, the estimated coefficient of, say, EATPD represents an estimate of the average incremental costs resulting from a percentage point increase in the proportion of total patients who are partially dependent on eating. By using the raw percentages in the cost function, the combination of debility levels and debility types is accounted for. None of the other indices provide this information. As emphasized throughout it is the type of debilities present in a patient or home that influence costs. Using the raw percentages in the cost function also accounts for the range of disabilities with which each home must contend. This property, lacking in each of the other indices examined, enables one to capture the costs associated with the degree of dispersion of debility types present in a home.
Finally note that the alternative case-mix measures examined add little if anything to model accuracy. Hence, further regression work will use only the straightforward 12 variable index which measures the percentages of patients in the six different ADL categories.

T-4.3.7 Assessing the Impact of Variable Groups

Beta coefficients were examined in a further effort to assess the impact of different variables on cost. For various groups of variables, the beta coefficients were summed.[1] The table below reports the summed beta values for the 1972 data using 10 groups of variables.

Table T-4.4

SUMMED BETA COEFFICIENTS

Variable group	Beta Coefficient Sums
Services Offered	0.1668
Services Provided	0.1182
Capital Sources	-0.019
Admissions/Patient Day	0.2785
Input Prices	0.4179
Urban-Rural	-0.0324
Patient Age	-0.0093
Patient Diagnosis	0.1564
Marital Status	0.0123
Activities of Daily Living	0.0946

[1] The beta coefficients are simply normalized regression coefficients and bear a straightforward relationship to the unnormalized regression coefficients, being equal to the unnormalized coefficient times the ratio of standard deviations for the independent (S_x) and dependent (S_y) variables:

$$b_x \cdot \frac{S_x}{S_y}$$

A beta coefficient of, say, .25 can be interpreted to mean that a one-standard deviation change in the independent variable leads to a .25 standard deviation change in average operating costs, the dependent variable. (It is interesting to note that the beta coefficient for the independent variable in a bivariate regression model is identically equal to the simple correlation between the two variables.)

Using the beta coefficient sums as measures of the impact of different groups of variables on costs, we find that the variables measuring input prices have the largest impact: a one standard deviation increase in each input price variable leads to a 0.42 increase in per diem operating costs, certeris paribus. Other groups of variables have markedly smaller impacts on costs. Case flow, as measured by the number of admissions per patient day, has the second largest impact on costs. A one standard deviation increase in this variable leads to a net 0.28 standard deviation increase in costs, certeris paribus.

The complete cost model fit here has a relatively high degree of multi-collinearity due to the large number of independent variables, many of which measure similar factors. In Table T-4.5 we report on attempts to assess the statistical significance of groups of variables. The F-tests applied are not as affected by multi-collinearity problems as are the individual t-tests used elsewhere in this paper. The variable groups are significant at the 0.05 level or better if the F-statistic is greater than or equal to the criterion value. Summarizing, the following groups of variables are significant at the 0.05 level or better: (1) average daily census; (2) average daily census and occupancy; (3) patients diagnoses; (4) activities of daily living, diagnosis, age and sex; (5) all services; (6) services offered; and (7) regulatory variables. As can be seen, three different measures of patient-mix were reported in the table along with a combined measure of patient mix. Surprisingly, only the group of variables measuring patients' diagnoses and the combined descriptor were found to be significant. The significance of the combined variable may also result from the overwhelming effects of the diagnoses descriptor. The measures of utilization were found to be insignificant when taken alone. When combined with measures of scale, however, they were found to be significant. Measures of scale, on the other hand were found to be significant in isolation, therefore, begging to question as to whether the significance of the combined variable resulted from the scale effect overwhelming the utilization effect. The reader should bear in mind that these results depend on the specification of the model within which the variable groups have been tested.

Table T-4.5

RESIDUAL F-TEST RESULTS

Variable Group	F-Statistic	Significance Criterion For 0.05 Level
(1) Occupancy (OC1-OC6)	.23	2.10
(2) Average Daily Census (AD1-AD6)	11.41	2.10
(3) Average Daily Census and Occupancy	6.70	1.76
(4) Activities of Daily Living (GROOMPD, GROOMD, DRESSPD, DRESSD, EATPD, EATD, WALKPD, WALKD, BEDPD, BEDD, CONTINPD, CONTIND)	1.29	1.76
(5) Patients' Diagnoses (ARTERY, MENTAL, STROKE, ACCIDENT, NUTRIT, OLD, RESPIRAT, DIGEST, NEOPLASM, HEART)	1.82	1.76
(6) Age and Sex (AGE, FEMALE)	0.48	3.00
(7) Activities of Daily Living, Diagnosis, Age and Sex	1.58	1.47
(8) Services - All Measured (VARIABLES LISTED IN 9 AND 10)	2.21	1.58
(9) Services Offered (OT, RT, REHAB, EDUC, PT, SPHT, COUNSEL)	2.77	2.02
(10) Services Provided (NASAL, BLOOD, ENEMA, CATH, BATH, BOWEL, OXYGEN, BAND, PULSE, HYPO, IRRIG, INTRAV, SUMSER, PONEB, PTHREEB)	1.49	1.76

Table T-4.5 (Cont'd.)

RESIDUAL F-TEST RESULTS

Variable Group	F-Statistic	Significance Criterion For 0.05 Level
(11) Regulatory Variables (ICF, ICF-SNF, PCTPRI, PCTCARE, PM1, PM2, CON, LSC, LIM, STAFFA, STAFFB, MA, UNIF, CB, OCCUPEN, NORTHEAST, CENTRAL, SOUTH)	8.50	1.80
(12) Payment Proportion (PCTPRI, PCTCARE)	.05	3.00
(13) Facility Variables (PROFIT, NONPROFIT, CHANGE, CHAIN, PCONTRACT)	.72	2.61

T-4.3.8 Outlier Analysis

This section examines the standardized residual terms associated with the complete model just discussed at length. The purpose is to demonstrate that there are no discernable patterns in the residuals, i.e., there is no evident correlation and/or heteroskedasticity among distrubances. Of special interest are the characteristics of facilities for which the estimated disturbance is positive and large in magnitude. These homes have been isolated and their general characteristics examined.

T-4.3.8.1 Characteristics of Predicted Observations

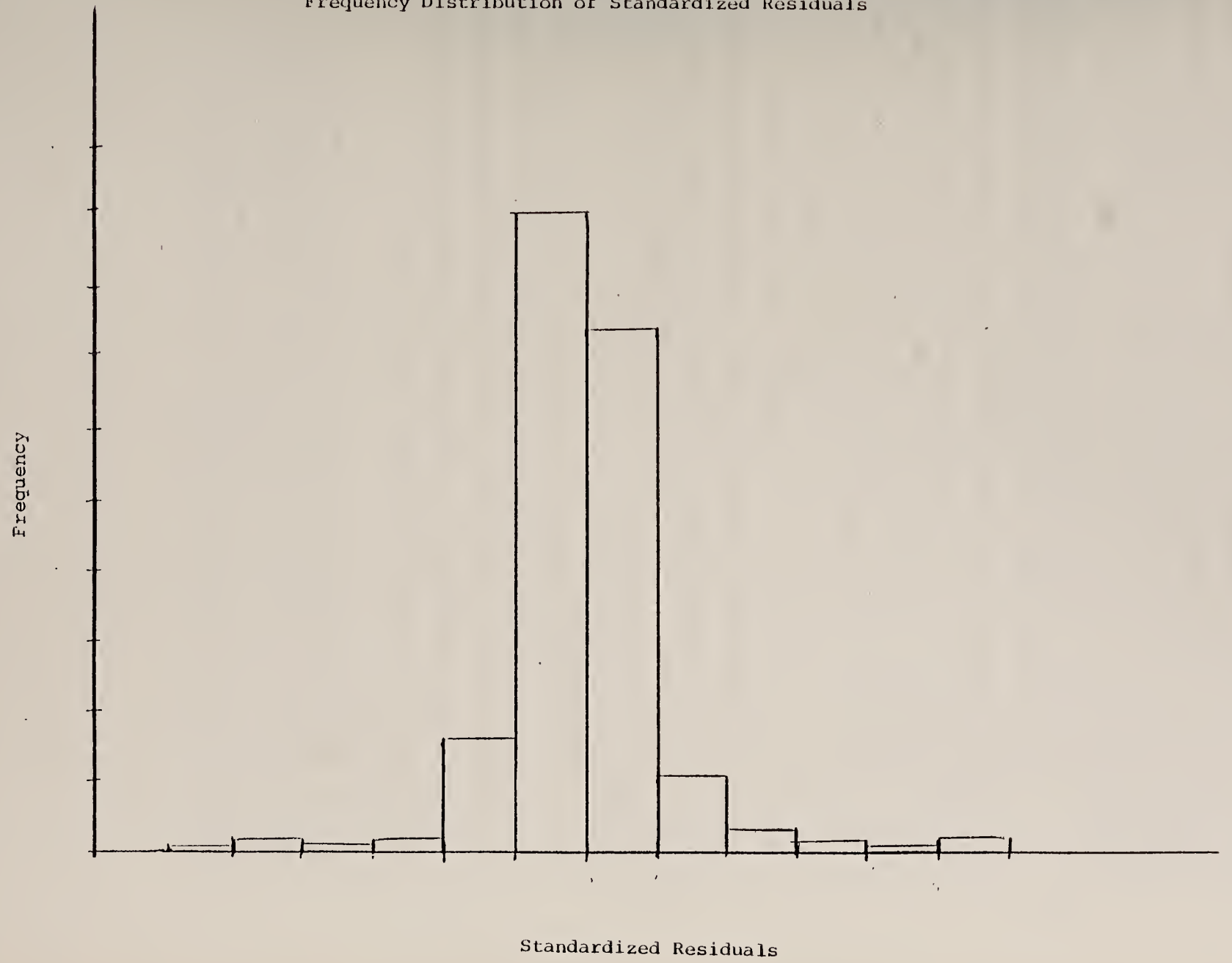
Inspecting patterns in the standardized residuals can uncover possible faults in the predictive model, specifically, a potential bias toward homes with particular characteristics. It is demonstrated here that there is no obvious bias in the model in terms of either overestimation and/or underestimation of the operating costs of certain types of homes.

Figure T-4.4 shows a histogram of the standandized residuals. The distribution is symmetric and appears normal.

To assess any implicit favorable or unfavorable bias present in the model, the basic characteristics of nursing homes were examined for twelve ranges of the standardized residuals. First, for each of the twelve ranges, means and standard deviations were obtained for the variables, average operating expenses, (ADC), bedsize of the home (BEDS), census of the home (CENSUS), and the occupancy rate (OCC). The results are reported in Table T-4.B-4, Appendix T-4.B. From inspection it is seen that for each of the four variables, there is considerable overlap in the distribution of each variable across the ranges of the residual. In particular, the means for the four variables in the very high and the very low range are within one standard deviation of one another indicating that in each case, we fail to reject the hypothesis that the means are equal. Thus homes whose costs are over or under-predicted cannot be characterized in terms of these variables.

Figure T-4.4

Frequency Distribution of Standardized Residuals



Second, for each of the twelve ranges, frequency distributions were obtained for the four discrete variables, urban-rural classification (SMSA), ownership type (OWNER), certification type (CERT), and region of the country (REGION). The results are summarized in Table T-4.B-6, Appendix T-4.B. Again, there appears to be no obvious characterization of nursing homes in each of the selected ranges of the residuals.

T-4.3.8.2 Correlation and/or Heteroskedasticity in the Residuals

In a further attempt to characterize those homes with large or small residuals, the estimated disturbances were examined for possible patterns of correlation and patterns of heteroskedasticity. The mean and standard error of the disturbances were estimated for residual subgroups based on urban-rural status, certification status, ownership type, and regional location. The results are reported in Table T-4.B-5, Appendix T-4.B. Inspection of the table reveals that by each criterion variable the residual means do not differ significantly from 0, and the standard error of the residuals by subgroup does not differ significantly from 1. These results support the adequacy of the assumptions made concerning the model, namely that the disturbance term of the model is normally and independently distributed with zero mean and constant variance.

T-4.4 Limited Model

This study is exploratory in nature. It has been designed to investigate a broad range of hypotheses about factors which may or may not affect operating costs. Thus, while all the explanatory variables in the "complete" model have theoretic justification, we have been rather generous in including a large number variables in the cost function.[1]

---

[1] Some of these variables are close proxies for other variables (e.g., Staff A and Staff B measure different types of staff input requirements); others have been included because they have been used elsewhere (e.g., patient sex) or because their inclusion is initially necessary to test a particular hypothesis (e.g., the variables depicting service intensity); finally, some have been included to control as much as possible for variables affecting the home environment (e.g., the regulatory variables).

The complete model, discussed above in Section T-4.3 provided a "good fit" of factors explaining average operating costs of nursing homes. While the interpretation of the coefficients (for the most part) followed *a priori* expectations, the large number of explanatory variables, many of which measured related factors, allowed for some uncertainty of interpretation, given the high potential for multi-collinearity. Moreover, their inclusion adds greatly to the complexity of interpretation. While factor analysis and other techniques are available for aggregating variables, such techniques obfuscate the impact of particular variables and are not appropriate for rate setting purposes where it is important to know whether particular variables impact on nursing home costs. Nevertheless, to present a rate setter with a list of over 90 variables and state that all need to be included in a rate formula would only add to the general skepticism of regulators and other non-economists about the utility of econometric results. Nor is it clear that these statistically insignificant variables add anything to the exploratory power of the model.

The discussion below reports on a somewhat *ad hoc* effort to alleviate these problems. The "limited" (or partial) model presented in Table T-4.6 includes only those variables with a F-statistic greater than or equal to one. This strategy corresponds to a rough "rule of thumb" among econometricians that coefficient estimates greater than their standard errors may be reliable. In lay-terms, a F-statistic equal to one means that the chance of a coefficient being zero is approximately 30 percent. There were only three exceptions to this rule. The average daily census variables (AD_i), the certification dummies, and the occupancy variables (OC^1) were forced to enter the regression despite their F-value. The reason here again was to make the results palatable to rate setters who typically consider the scale of the nursing home and its occupancy rate as fundamental ingredients to rate determinations. Conversations with persons knowledgeable about the rate setting process suggested that the inclusion of these variables was necessary, if only to show their insignificance.

Table T-4.6: DETERMINANTS OF AVERAGE OPERATING COSTS: THE LIMITED COST MODEL; NATIONAL NCHS DATA

Explanatory Variable	Estimated Coefficient	Standard Error
Output: Scale		
CONSTANT	23.2974	
AD1	-0.9497xxx	0.2084
AD2	-0.0692	0.0428
AD3	-0.0318	0.0094
AD4	0.0150	0.0138
AD5	-0.0243	0.0174
AD6	0.0123	0.0064
OC1	-0.0186	0.0356
OC2	-0.1616	0.1998
OC3	0.1185	0.2139
OC4	-0.2444	0.2519
OC5	0.0459	0.1474
OC6	-0.0630	0.1635
Diagnosis		
ARTERY	0.0232xx	0.0086
MENTAL	-0.0188	0.0114
STROKE	0.0202	0.0129
ACCIDENT	0.0262	0.0189
NUTRIT	-0.0289	0.0215
Mental Status		
DEPRES	0.0164xx	0.0070
ABUSE	-0.0134	0.0105
SLEEP	-0.0091	0.0082
Activities of Daily Living		
GROOMPD	0.0141x	0.0072
DRESSD	-0.0123	0.0088
EATPD	0.0152	0.0104
EATD	0.0640	0.0485
BEDD	0.0198	0.0188
CONTINPD	-0.0033x	0.0166
Output Services		
OT	0.928xx	0.3760
RT	0.4480	0.3474
EDUC	0.4493	0.3120
PT	0.9940xx	0.4160
BLOOD	0.0090x	0.0053
ENEMA	0.0190x	0.0090
CATH	0.0020	0.0200
PULSE	0.0054	0.0049
SUMSER	0.0743xxx	0.0254
PONEB	0.0092	0.0097

Table T-4.6: DETERMINANTS OF AVERAGE OPERATING COSTS: THE LIMITED COST MODEL; (Cont'd.) NATIONAL NCHS DATA

Explanatory Variable	Estimated Coefficient	Standard Error
Output: Case Flow		
ADMISPD	434.8556xxx	39.3442
Input Prices		
RETAILWG	0.9653xxx	0.3539
NURSEWG	2.8168xxx	0.1862
SMSA	-0.4428	0.3476
Regulatory Variables		
ICF	-2.2750xxx	0.4443
ICF-SNF	-0.5303	0.3718
PTCARE	0.0332x	0.0176
PM1	-1.3423xxx	0.4130
PM2	-1.2972xxx	0.4894
LSC	0.8645xx	0.3605
STAFFA	1.5996xxx	0.5926
UNIF	0.8586x	0.4968
OCCUPEN	0.9144x	0.4803
NORTHEAST	3.3175xxx	0.5652
CENTRAL	1.3466xx	0.5945
SOUTH	2.2664xxx	0.5330
Facility Characteristics		
PROFIT	-2.4018xxx	0.6283
NONPROFIT	-0.6776	0.6728
PCONTRACT	0.5473	0.0352

R^2	0.5791
Adjusted R^2	0.5575
F-ratio	26.7946
N	1127
Standard Error	4.744

x significant at .10 level
xx significant at .05 level
xxx significant at .01 level

T-4.4.1 Results of the "Limited" Model

Adoption of this convention resulted in the elimination of 36 variables, a decrease of 40 percent from the 91 variables in the "complete model."[1] Nevertheless, the limited model shown in Table T-4.6, explained virtually the same fraction of the variation in average operating costs. (The R^2 was 0.5791 versus 0.5847). Moreover, the adjusted R^2, which corrects for degrees of freedom, increased from 0.5482 to 0.5575. Finally, those variables present in the limited equation retained coefficients virtually identical--both in absolute magnitude and statistical significance--to previous estimates in the "complete" equation.

The limited model produced the following results. In terms of economies of scale, there were statistically significant returns to average daily census for $0<ADC\leq100$; insignificant coefficients for $100<ADC\leq200$; and a positive and significant coefficient for $ADC>200$. Again, as in the complete model, there were no returns to increasing utilization; all occupancy variables were insignificant. Hence, we still conclude that the optimal size of a nursing home is in excess of 100 beds, since beyond that range, the curve is virtually flat. The estimated cost per admission has risen slightly to $434, and maintains its high level of significance.

With regard to patient characteristics, statistically significant The percentage of patients with a hardening of the arteries and the percentage of patients depressed were statistically significant. While the estimated coefficients were small, they would have a relatively large cost impact only if a significant percentage of the patients in the home were in these categories. In terms of the activities of daily living measures of ability, patients partially incontinent or needing assistance in dressing reduce cost dressing reduce cost slightly; patients dependent in grooming, eating, or requiring assistance in moving in or out of bed increase costs slightly. Neither age nor sex enter the regression. Thus, somewhat more of the patient

---

[1] For the most part, the deleted variables were the patient descriptors (the diagnosis and ADL variables) and the service intensity variables; some of the services offered and regulatory variables also were deleted.

characteristics enter into the policy model. Nevertheless the economic consequence of these patient characteristics still is of a far lower order of magnitude than the other measures of output--or of the measures of input prices, regulation, and facility characteristics. The logic here followed that of Section 4.1.2.3. The coefficients of all patient characteristics which entered the model (and hence had on F>1) were multiplied by the mean percentage of patients having these characteristics and summed over characteristics. The result is that the estimated maximum average effect of patient characteristics is around $1.50, or less than 11 percent of the mean of average operating costs of the homes in the data. If we consider only those variables with a coefficient statistically significant at the 0.10 level, the resulting cost impact is a half, $0.77, or less than 6 percent of average operating costs of homes in the data.

Turning to the service variables, the policy and complete models again give comparable results. Offering either occupational or physical therapy increases per diem costs by over $0.90 per service; both are statistically significant. Offering educational programs or rehabilitation therapy are marginally significant variables with per diem cost impacts of over $0.40; note that where in the complete model the rehabilitation variable was negative, its sign here has reversed to a more intuitive, positive effect. Looking at the service intensity variables reveals that enema services retain their statistical significance; in addition, blood pressure reading has become statistically significant, and catheterization has become marginally significant. As before SUMSER too retains its statistical significance, and the bed size variables are still insignificant. As in the complete model, the dollar magnitude of the coefficients of service intensity variables is far less than the cost impact of offering services.

The input price variables enter in at the same high levels of economic and statistical significance. The coefficient of RETAILWG is $0.97 and the coefficient of NURSEWG is $2.82.

With regard to the regulatory variables, as before, ICF homes have per diem costs of about $2.20 less, and ICF-SNF homes have per diem costs of about $0.50 less than purely SNF homes. Again, ICF was highly significant and ICF-SNF was insignificant. With regard to payment source, as before, Medicare patients have higher ($0.03) per diem costs than other patients.

Turning to the impact of specific regulations, it is still the case that homes on prospective reimbursement systems have higher per diem costs (on the order of $1.30) than homes on cost-related or flat rate systems. Again as before, absolute staff, life safety code, and occupancy requirements retain high levels of statistical significance (with cost impact of $1.60, $0.86, and $0.91 respectively). Where uniform chart of accounts was of only marginal significance in the complete model, it is now significant at a .10 level (with a coefficient of $0.86).

With regard to region, homes in the Northeast region had the highest costs, followed by homes in the Central and then the South regions.

The final set of variables were the additional facility characteristics. For-profit homes still have significantly lower costs ($2.40) than do government homes; non-profit status was still statistically insignificant. The conclusion remains that patients in for-profit facilities have lower costs, given the measures of output allowed by the data, than do non-profit/government homes. Finally, the extent of contracting out for services is still insignificant although the positive coefficient suggests that homes with more contracting have slightly higher costs.

In conclusion, it is clear that both the complete model and the limited model yield very comparable results. While a few more of the patient characteristics enter into the policy model (for the most part with marginal significance)--no doubt because the multi-collinearity problem has been reduced--the remainder of the significant variables in the policy model also enter the complete model. In terms of coefficient changes, the "limited" model coefficients are, in general, between 0-10 percent higher than in the complete model; the exceptions here are the regional variables (which decline by 10-20 percent) and ownership status (the for-profit coefficient declines by five percent). The conformance of the results with economic theory as well as their robustness is notable.

T-4.4.2 Which One: The Complete Model or the Limited Model?

The next issue is to consider which model is "most" appropriate for rate setters. All variables in the complete model are based on economic theory; while multi-collinearity is clearly present in the complete model, no doubt it is still present in the limited model results too. However in both models, the same variables retain statistical significance and maintain their relative economic magnitude.

The purpose of this national analysis is to gain an appreciation of the general determinants of nursing home operating costs--to find out what variables are significant and what the magnitude of their impact is; it is not to develop specific parameters to be "plugged into" a rate formula. For purposes of rate setting, what one learns from both models is comparable. Moreover, the state specific analyses reported elsewhere[1] reveal that the specific coefficients of the cost functions differ across the states, although their general impact is similar. Thus, pragmatically, there is no reason to choose either the complete or the limited model for describing the national data; both only approximate what is happening at the state level. Econometric theory urges the analyst to select the model which most closely follows economic theory; however, it also suggests that the analyst weigh the associated costs of complexity and inaccuracy in each model.

Since the costs of inaccuracy may be large at a policy level, we conclude that the complete model should be the basis for further discussions of average operating costs with respect to national costs.

T-4.5 Summary

The analyses reported here demonstrate the applicability of economic cost modeling techniques to the nursing home industry. While the results do not directly translate into state-specific reimbursement formulae, they are useful in capturing the determinants of variation in per diem operating costs among a national sample of nursing homes. The conclusions developed here summarize the results presented above in the context of results found elsewhere. Comparison of our results with the other studies underlines the importance of correct specification of nursing home cost models.

---

[1] Technical Papers #5, #6, and #7 to this study.

Optimal Size

Results here suggest that there are slight economies of scale up to an average daily census (or bed size) of 40 patients. Beyond that point, average operating costs appear relatively insensitive to the number of patients.

Evidence in other studies on the effect of the size of nursing homes on costs is mixed, and may be attributable to incomplete specification of the variables and a quadratic rather than a "grafting" approach to size specification. The Moreland Commission (1976) found that size (measured by square footage of the institution) was the key determinant of cost. However, they did not control for either patient characteristics or services provided. Moreover, it is not clear that a measure of capital intensity should be entered into an operating cost equation; no measures of aveage daily census or bed size were reported. Ruchlin (1972) found that size (measured by patient days) was not significantly related to average cost; the estimated coefficients were statistically insignificant and produced the counter-intuitive result that costs increased at first, and then decreased with increases in patient days--a cost curve with an inverted U shape.[1] Again, services offered were omitted, as well as patient characteristics except for payment mechanism. Skinner and Yett (1970), too, found a disparate set of results. Using beds as a measure of output, they too found an inverted U shape, though not statistically significant. With a patient day measure of output, they found that costs increased monotinically. Using a model which included both patient stock (number of cases) and patient flow (total bed days), the estimated cost curve had a very flat, U shape, with a mimimum around 85 beds. In general, they seem to conclude that the cost curve can best be described as having a very shallow "saucer" shape, a result in line with those found here. Skinner and Yett pointed out the possibilities for mis-

[1] Cost equations results for the individual components of operating cost (e.g., food) did produce U shaped cost curves.

specification and multi-collinearity in their results;[1] moreover, they too omitted services offered, as well as input prices. The AMS (1976) study concluded that bed size positively affected costs; their final equation entered only "bedsize greater than 60"; this variable raised costs by $1.50. Because AMS defined per patient cost in terms of the rate charged to private patients, the relation of this result to the others is not clear.

Occupancy

Results here found no occupancy effect. Elsewhere, Ruchlin as well as Skinner and Yett entered the occupancy rate in a linear fashion and found that it was negatively related to cost. In their final result, AMS entered only the single term "occupancy rate over 93 percent." The difference in results may be due to either omitted variables in the other studies, or to state specific data of Ruchlin (Massachusetts) and Skinner and Yett (California).

Patient Mix

It is widely believed that patient mix affects nursing home costs. Little evidence was found here for such effects. However, as pointed out previously, the data were such that the impact of patient mix was underestimated. In contrast to Skinner and Yetts' results that the average age negatively affects costs and percent female positively affects costs, neither variable here was of economic or statistical significance.[2]

With regard to patient health status results here found that per diem cost was, for the most part, unrelated to diagnosis, mental condition, or measures of debility--after controlling for services offered. Of course, if all the patients in a home did have those characteristics which were statistically significant or marginally (here, take to be F>1), the cost impact could be large. However, an analysis of the average effect of patient characteristics suggested that the maximum effect was on the order

---

[1] In the data used for this paper the correlation between patient days and patient days squared is .909.

[2] Neither Ruchlin nor AMS report age variables; AMS reported sex as insignificant.

of 6-11 percent of average per diem costs. This result is in contrast with the Skinner and Yett finding that partially debilitated patients cost more than either independent or dependent patients[1]; however, the Moreland Commission found that ambulatory/non-ambulatory status was not a significant factor and concluded that cost variation had nothing to do with patient need.

Services

A major findings was the impact of services offered on cost. That the intensity of service use was not significant suggests that costs vary with the absolute availability of services, rather than their relative use. Only the AMS study allowed explicitly for services offered; these variables were related to their proxy measure of cost, private price.

Input Prices

As found elsewhere, input prices were positively related to average operating costs.

Regulations

Little is known about the effect of regulations on nursing homes, aside from the JWK (1976) simulation of SNF/ICF operations standards. Also, Ruchlin's study of Massachusetts' homes over the period 1965-1969 found that alterations in reimbursement procedures were related to costs.

Results here suggest that there is a definite relationship between regulations and nursing home costs. However, the relationship is complex and warrants more investigation as is seen in the coefficients of the reimbursement type variables. It was suggested above that the positive association of prospective rate systems (as opposed to cost-based or flat rate systems) with higher costs may not be causal; while prospective rate systems may cause higher costs--for example, if facility operators require a risk premium, or if administrative costs are higher--results here may also reflect a historical phenomenon--high cost states are likely to be most responsive to new techniques to contain these costs. Other regulatory variables were also

---

[1] AMS found that while partially debilitated patients cost more than independent patients, totally debilitated patients cost the most.

associated with costs differentials; staffing and occupancy requirements were positively related to costs, while payment limits may reduce costs. Similarly, licensure (SNF/ICF) status (which may also be considered a proxy for services offered and patient condition) also affects costs.

Finally, it is widely believed that private patients pay higher rates for nursing home care than do public patients. This is commonly transformed into an assumption that private patients cost more than public patients. However, this result may only reflect the residence of private patients in homes offering more services, and hence does not mean that, all other things equal, private patients cost more than public patients. Both Skinner and Yett as well as Ruchlin found that private patients cost more than public patients. However, neither included measures of services offered--a result comparable to our preliminary analyses which excluded the service variables. After including measures of services offered, it was found that private and Medicaid patients cost the same, while Medicare patients cost somewhat more.

APPENDIX T-4.A
DISCUSSION OF DATA

T-4.A.1 Data

Data used in the national nursing home analyses, Technical Paper #4, come from three types of sources. Data on the nursing homes, including facility and patient characteristics, come from the confidential version of 1973-74 National Nursing Home Survey conducted by the National Center for Health Statistics, and is referred to as the NCHS data. Data on the county in which the nursing home was located come from the Area Resource File, referred to as the ARF. Data on state regulatory practices come from a variety of surveys, as discussed below.

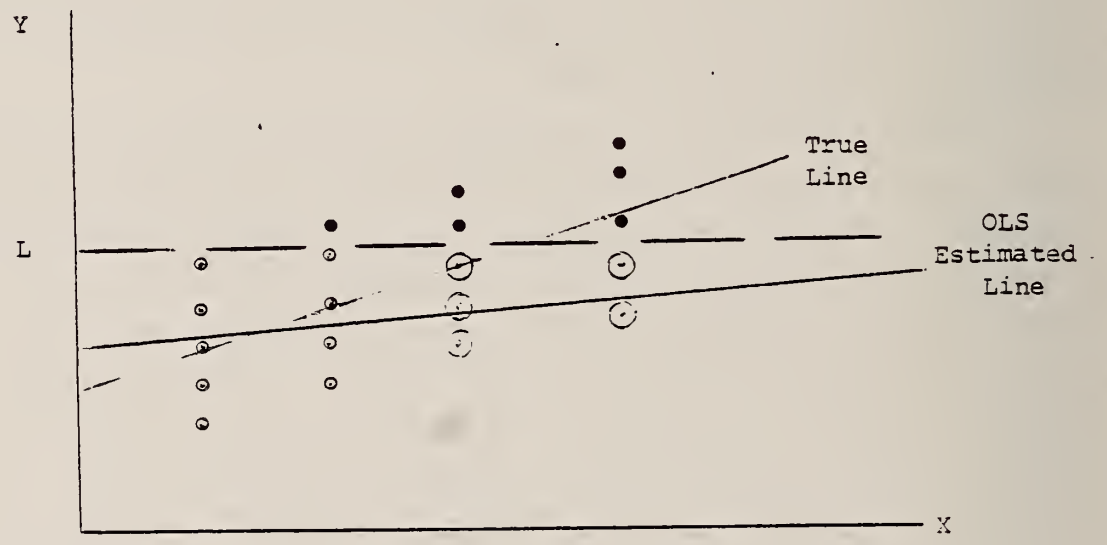
The NCHS Data

The 1973-74 National Nursing Home Survey was a sample survey conducted between August 1973 and April 1974 by the National Center for Health Statistics. The universe used to draw the sample was the 1971 Master Facility Inventory of all nursing facilities in the continental United States. To be included in the NCHS sample, the facility had to qualify as either a nursing care home or a personal care home with nursing, thus excluding facilities offering only personal or domiciliary care. Also excluded were facilities that opened in 1973 or whose level of care was upgraded from personal care to nursing care in 1973.

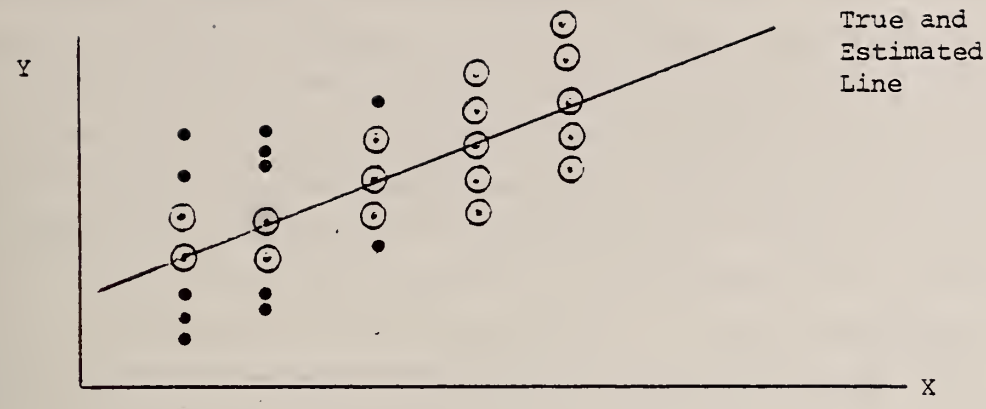
Sampling was based on a stratified two-stage probability design. The first stage selected facilities and the second stage selected residents and employees in the sampled facilities. Prior to sampling, the universe of homes (17,685) was stratified by Medicare, Medicaid certification, bed capacity, ownership, geographic region, state, and county. The first stage sample was then selected randomly from each strata. The second stage sample selected a maximum of 10 residents and 14 staff per facility. Data were collected on facility characeristics, expenses, residents and staff. The survey met NCHS standards that reliable estimates are those having a relative standard error of less than 25 percent. Further detail on the NCHS data can be found in National Center for Health Statistics (1977).

Some comments on the applicability of using a least squares regression technique on data not obtained by simple random sampling are appropriate.

The NCHS nursing home data were obtained as a stratified random sample with variable sampling fraction between strata. This gives a sample that is not representative in the aggregate. Even so, it is clear that least-squares estimation from this data will give unbiased parameter estimates--assuming the model is correctly specified. The classical linear regression model does not require a proportionate (i.e., simple random) sample. This point can be illustrated both algebraically and graphically. Below we choose the more heuristic, graphic approach. Consider an experiment in which we would want to draw population inferences about earnings using a sample of only low income families. Thus, the sample is effectively "truncated" at L, a dependent variable level, as shown below. If the points in the population are indicated by dots and the sampled ones by circled dots, we get the following configurations. Note that the line estimated by OLS will be biased.



Alternatively, suppose that the sample proportions do not reflect the actual distribution along the independent variable dimension, X. In this case, the shape of the curve is not affected, as shown below. The only statistic which can be influenced by sample methodology is the variance of the estimated coefficient in various ranges of this curve. The literature is surprisingly sparse on this relationship.



Bias in sampling methodology arises only when the sample is stratified on the basis of the dependent, not the independent variables. For example, this is a problem which pervades results from the New Jersey income maintenance experiment where income was used, de facto, to restrict participation. Here, the problem of statistical bias was serious. To deduce whether this problem is embodied in the NCHS data, we evaluated the dimensions along which the sample was stratified. In no instance was a dependent variable used as a partitioning statistic.

Nevertheless, it is also clear that the efficiency of the estimates is affected by the manner in which the sample is drawn. This problem pervades much econometric work--and yet, so far as we have been able to determine, it has not been explicitly addressed in the existing literature. It seems that the answer is ambiguous, depending upon the specific empirical context of the analysis. Thus, a simple random sample of nursing homes may or may not give greater efficiency than the NCHS sample--holding the sample size constant.

One last comment on the NCHS data is in order. Expense data were reported by the facility. While the questionnaire asked about costs with reference to particular cost components, e.g., wages, laundry, maintenance, etc., there was no common convention of allocation, except according to respondent interpretation. Thus, the expense data do not conform to a uniform chart of accounts.

The ARF Data

Data on the county in which the nursing home was located come from the Area Resource File (ARF) assembled by the Bureau of Health Manpower, DHEW from a wide range of sources. The ARF includes data on all counties and Health Systems Agencies in the nation, including information on local population characteristics, health manpower, health facilities, and health expenditures. Because the confidential version of the NCHS data contained county identifiers, it was possible to uniquely identify nursing homes with appropriate county data from the ARF.

Regulatory Data

No single source of data was available here. Variables describing state regulatory policies based on Applied Management Sciences (April 1976), Social Rehabilitative Services (1973) and Urban Institute (N.D.), as validated and supplementeed by other sources. The data was combined to produce a description of the reimbursement and regulatory systems in the states appropriate to 1973-74.

The AMS study provided a non-systematic narrative discussion of the reimbursement and regulatory systems in the states as of January 1, 1975. Because the AMS data was for a later point in time than the NCHS data, as well as being incomplete and unverified, it was supplemented by surveys conducted by the Urban Institute, the Government Accounting Office, Battelle Health Care Study Center, and DHEW Social Rehabilitative Services (SRS).

The Urban Insitute (N.D.) provided copies of information gathered in a phone survey of the 50 states on their reimbursement policies, rates, eligibility requirements, and certificate of need requirements for the period 1963-73.

The General Accounting Office (1972) summarized state reimbursment policies for 1972. Although the information is not exactly comparable to AMS and Urban Institute data, it does provide reimbursement specific information, e.g., whether the reimbursement system is based on actual or reasonable cost, negotiated, etc.

The Battelle data (N.D.) includes information on the reimbursement procedures in 21 states for 1974.

The SRS data (1973) is based on a questionnaire compiled by the Social and Rehabilitation Service in 1973 and 1974 on the 1973 Medicaid reimbursement rates and policices.

All the above sources face a common set of problems. First, they were not designed for estimating the effect of reimbursement practices on nursing home costs. Many of the variables of interest to this study are not included: e.g., what grouping criteria were used? What cost variables had ceilings and what were the levels of the ceilings? What was the required occupancy level? What input standards were in effect? etc. Second, the responses to questions on state reimbursement practices relied on state interpretations to open ended questions rather than a choice among a rigorous formulation of possible answers.[1] That a state says it had a prospective reimbursement system still leaves much ambiguity since, as discussed in Chapter 2, "prospective reimbursement" can mean a multitude of things. Third, with some exceptions, notably the SRS and UI data, the information are for years other than 1972. The first problem reduces the richness of the anlaysis which can be undertaken and allows the possibility that some state regulatory differences have not been controlled. While the second problem cannot be dealt with rigorously, use of a variety of ad hoc procedures resulted in a reasonably accurate set of variables. An initial attempt to resolve differences between sources involved some of the states (there were differences in approximately 4-8 states in terms of payment mechanisms between any two of the data sources). After several exploratory conversations it became clear that it was both difficult (if not impossible) to locate individuals knowledgeable about 1972, and that reliable statements were not possible.[2] As a result, separate regressions were estimated, initially, with data as specified in each of the sources. In all cases, results of using the alternative data sets were virtually identical for both the regulatory coefficients and the coefficients of other variables in the regression. Because

---

[1] For example, the Urban Institute asked "what reimbursement system was used by the state in each of the years 1966-1973?" While a range of alternatives were posited for selection (e.g., reasonable cost, cost related, prospective reasonable cost, etc.) explicit definitions were not provided.

[2] Because memories and records do not appear to span the period between 1972 and 1977, the government Project Officer for this study directed that we stop such efforts and rely on existing data.

of the robustness of these measures, we felt comfortable in relying on the two sources with the most comprehensive set of reimbursement data, the SRS and AMS data. These data were verified, and modified (if necessary) with information on state-specific regulatory practices found in the other data bases.

With regard to resolving differences in timing of the reimbursement information, the following decision format was adopted for resolving differences among the sources. If all sources agreed that the rates were cost based, prospective, or flat/point system, the common answer was accepted. If the UI, AMS, and SRS data agreed, but the GAO data disagreed, we accepted the UI/AMS/SRS data since they are for the adjacent or correct year. If there was disagreement between the UI, AMS, and SRS data, we accept the AMS/SRS data since the data here are supplemented by more descriptive information appropriate for evaluating the suitability of the answer. A final check was to review inconsistencies with other researchers using these same sources of data. Thus, it was possible to merge the answers into a unified format.

Variables defining the reimbursement and regulatory system are described below. The data are summarized in Table T-4.A-1. For reference, "prospective rate" "flat rate" and "restrospective rate," are defined as:

> Prospective rate: a rate-setting mechanism in which the rate for an individual facility for a specific period is determined prior to the period in question with no automatic retrospective adjustment.
>
> Flat rate: a rate-setting mechanism in which the rate for a specific period is set at one level for a group of facilities. The rate may be negotiated, legislated, or established by a point system.
>
> Cost-based: a rate-setting mechanism in which the rate for a specific period was determined during or after the period in question. Typical procedures here are to follow the Medicare guidelines of paying all "allowable costs" or alternatively to pay "cost-related," or "actual-cost" rates.

Table T-4.A-1: STATE REGULATORY AND REIMBURSEMENT SYSTEMS

STATE	PMI[1]	CB[2]	LIM[3]	OCCUPEN[4]	UNIF[5]	LSC[6]	MA[7]	CON[8]	STAFFA[9]	STAFFB[10]	RATE[11]	PM2[12]	PM3[13]	PM4[14]	PM[15]
Alabama	1	1	1	1	0	1	0	1	0	0.1	12.5	1	1	1	1
Arkansas	3	0	1	0	0	0	0	0	0	0.15	13.9	4	2	2	2
California	1	0	1	1	0	0	1	1	0	0.0	15.0	1	1	1	1
Colorado	1	1	1	1	0	0	0	0	0	0.25	12.5	3	4	3	3
Connecticut	1	0	1	0	0	1	0	1	0	0.03	17.5	1	1	3	1
Delaware	2	1	1	0	0	1	0	0	0	0.31	17.0	1	1	1	1
Florida	2	1	1	0	0	0	0	0	0	0.03	11.7	3	3	2	2
Georgia	2	1	0	1	0	0	0	0	0	0.25	12.5	4	1	1	1
Hawaii	2	1	1	0	1	0	0	0	1	0.0	25.0	4	NA	1	1
Idaho	2	1	1	1	1	1	1	0	1	0.0	11.5	1	1	1	1
Indiana	1	0	1	1	0	0	0	0	1	0.0	17.5	4	4	3	3
Illinois	3	1	1	0	0	1	1	0	1	0.0	13.8	4	3	2	2
Iowa	2	1	1	0	0	0	0	0	0	0.21	17.5	1	1	1	1
Kansas	2	1	1	0	0	0	0	0	1	0.0	12.5	1	3	3	1
Kentucky	2	1	1	0	0	0	1	1	0	0.0	18.0	3	1	1	1
Louisiana	3	0	0	0	0	0	1	0	0	0.28	10.0	1	2	2	2
Maine	2	1	0	1	0	1	0	0	0	0.31	26.0	2	2	2	2
Maryland	2	1	1	0	0	1	0	1	0	0.25	16.5	1	1	1	1
Massachusetts	2	0	0	1	1	1	1	1	0	0.08	17.5	1	1	1	1
Michigan	2	1	0	1	1	1	0	0	0	0.13	18.0	2	1	1	3
Minnesota	1	1	1	1	1	0	0	1	0	0.25	13.5	1	1	1	1

1) 1 = prospective rate; 2 = retrospective rate; 3 = flat rate.

2) 1 = individual facility; 0 = grouping.

3) 1 = limits on total payments or specific categories.

4) 1 = occupancy penalties; 0 = otherwise.

5) 1 = uniform chart of accounts required; 0 = otherwise.

6) 1 = life safety codes in effect; 0 = otherwise.

7) 1 = medical assessments required prior to entry; 0 = otherwise.

8) 7 = certificate of need in effect; 0 = otherwise.

9) 1 = RN or LPN requirement; 0 = otherwise.

10) number of nurses required per patient day, day shift.

11) median operating cost rate reimbursement by state Medicaid program for (SNFs) ($s).

12) 1 = reasonable cost; 2 = negotiated rate; 3 = actual cost; 4 = flat rate.

13) 1 = reasonable cost or cost related; 2 = negotiated; 3 = flat rate; 4 = prospective rate.

14/15) 1 = cost related; 2 = flat rate; 3 = prospective.

See text for sources.

Table T-4.A-1: STATE REGULATORY AND REIMBURSEMENT SYSTEMS (Cont'd.)

STATE	PM1[1]	CB[2]	LIM[3]	OCCUPEN[4]	UNIF[5]	LSC[6]	MA[7]	CON[8]	STAFFA[9]	STAFFB[10]	RATE[11]	PM2[12]	PM3[13]	PM4[14]	PM[15]
Mississippi	1	0	0	1	0	1	1	0	0	0.0	12.5	3	1	1	1
Missouri	1	1	0	1	1	0	0	0	0	0.0	12.5	2	2	3	3
Montana	2	1	0	0	1	1	0	0	0	0.0	12.5	1	1	1	1
Nebraska	3	0	1	0	0	1	0	0	0	0.0	12.5	2	2	3	2
Nevada	2	1	1	0	1	0	0	0	0	0.0	22.5	2	NA	3	1
New Hampshire	2	1	1	0	0	0	0	1	0	0.0	22.5	1	1	1	1
New Jersey	1	1	1	1	1	1	0	1	0	0.31	21.0	1	1	1	1
New Mexico	1	1	1	0	0	1	0	0	0	0.0	15.0	3	1	1	1
New York	1	0	1	1	1	0	0	1	0	0.31	25.0	3	4	3	3
North Carolina	2	1	1	1	0	1	0	0	1	0.0	18.5	1	1	1	1
North Dakota	1	1	0	0	0	0	1	1	0	0.38	11.2	1	4	1	3
Ohio	1	1	1	0	1	0	0	0	0	0.04	15.0	4	3	2	3
Oklahoma	3	0	0	0	0	0	0	0	0	0.1	10.9	2	3	2	2
Oregon	1	1	1	1	1	1	0	0	0	0.25	10.3	1	1	3	1
Pennsylvania	3	0	1	0	0	1	0	1	0	0.31	12.5	1	3	2	2
Rhode Island	2	1	1	1	1	0	0	0	0	0.0	15.0	1	NA	1	1
South Carolina	1	1	1	0	0	0	1	1	1	0.0	16.1	2	3	2	2
South Dakota	2	1	1	1	0	0	0	0	0	0.2	12.5	3	1	1	1
Tennessee	2	0	1	0	0	1	0	1	0	0.0	22.5	1	1	1	1
Texas	3	0	1	0	0	0	0	0	0	0.03	14.1	2	3	2	2
Utah	3	0	1	0	0	1	0	0	1	0.0	12.3	2	3	2	2
Vermont	2	1	1	1	1	0	0	0	0	0.0	25.0	1	NA	1	1
Virginia	2	1	0	0	0	1	1	1	0	0.0	17.5	1	1	1	1
Washington	1	1	1	0	0	1	1	1	0	0.0	11.0	4	4	3	3
West Virginia	2	1	1	0	0	0	0	0	0	0.0	17.5	1	1	1	3
Wisconsin	2	1	1	0	1	1	0	0	0	0.25	15.0	2	1	3	1
Wyoming	1	0	1	0	1	1	1	1	0	0.28	12.0	2	4	2	3

1) 1 = prospective rate; 2 = retrospective rate; 3 = flat rate.

2) 1 = individual facility; 0 = grouping.

3) 1 = limits on total payments or specific categories.

4) 1 = occupancy penalties; 0 = otherwise.

5) 1 = uniform chart of accounts required; 0 = otherwise.

6) 1 = life safety codes in effect; 0 = otherwise.

7) 1 = medical assessments required prior to entry; 0 = otherwise.

8) 7 = certificate of need in effect; 0 = otherwise.

9) 1 = RN or LPN requirement; 0 = otherwise.

10) number of nurses required per patient day, day shift.

11) median operating cost rate reimbursement by state Medicaid program for (SNFs) ($s).

12) 1 = reasonable cost; 2 = negotiated rate; 3 = actual cost; 4 = flat rate.

13) 1 = reasonable cost or cost related; 2 = negotiated; 3 = flat rate; 4 = prospective rate.

14/15) 1 = cost related; 2 = flat rate; 3 = prospective.

See text for sources.

(1) PM1. Derived from the AMS data, this variable describes the payment mechanism: 1 = prospective rate; 2 = cost based, either actual, reasonable, negotiated or cost-related cost; 3 = flat rate, point system, or set by legislature.

(2) LIM. Derived from the AMS data, this variable describes whether the state specifies either a limit on total per patient payments per unit of time and/or limits on specified categories of cost such as food, laundry, administrative salaries, etc. If yes, then LIM = 1.

(3) OCCUPEN. Derived from the AMS data, this variable describes whether the state specifies that the reimbursement rate be associated with prescribed occupancy rates or if there are explicit occupancy penalties if the occupancy rate falls below a prescribed level. If yes, then OCCUPEN = 1.

(4) UNIF. Derived from the AMS data, this variable describes whether the state requires that nursing homes maintain a uniform chart of accounts. If yes, then UNIFAMS = 1.

(5) LSC. Derived from the Urban Institute data, this variable describes whether life safety codes were in effect for nursing homes. If yes, then LSC = 1.

(6) MA. Derived from the Urban Institute data, this variable describes whether medical assessments were required for patients prior to admission in the nursing home. If yes, then MA = 1.

(7) CON. Derived from the Urban Institute data, this variable describes whether certificate of need homes were in effect for nursing homes. If yes, then CON = 1.

(8) STAFFA. Derived from the Urban Institute data, this variable indicates that an RN or LPN be present irrespective of the size of the home. If yes, then STAFFA = 1.

(9) STAFFB. Derived from the Urban Institute data, this variable indicates the staffing level, in terms of the number of nurses required per patient day (ad justed for the SNS and ICF patient composition of the home) by state standards.

(10) RATE. This variable, derived from the SRS data, indicates the median operating cost rate reimbursed by the state Medicaid agency (for SNFs).

(11) PM2. This variable, derived from the GAO study, describes the payment mechanism: 1 = reasonable cost; 2 = negotiated rate; 3 = actual cost; and 4 = flat rate.

(12) PM3. This variable, derived from the Urban Institute survey, describes the payment mechanism: 1 = reasonable cost or cost related; 2 = negotiated; 3 = flat rate, legislated or point system; and 4 = prospective.

(13) PM4. This variable derived from the SRS survey, describes the payment mechanism: 1 = prospective rate, 2 = cost based; and 3 = flat rate.

(14) PM. This variable is the synthesized description of the payment mechanism: 1 = prospective rate; 2 = cost-based rate; and 3 = flat rate.

APPENDIX T-4.B

SUPPLEMENTAL TABLES

T-4.B.1 Introduction

The following tables were referred to in Technical Paper #4, which discussed the cost analysis of the national NCHS sample of nursing homes.

Table T-4.B-1: Cost Equations by Certification Level

Table T-4.B-2: Nursing Staff Hours Per Patient-Day Needed to Care for Selected Patient Types

Table T-4.B-3: Nursing Hours Per Patient-Day Needed to Care for Patients Dependent on Activities of Daily Living

Table T-4.B-4: Residual Analysis of Complete Model: Bias

Table T-4.B-5: Residual Analysis of Complete Model: Heteroskedasticity

Table T-4.B-6: Frequency Distribution of Selected Discrete Variables with the Standardized Residuals

Table T-4.B-7: Log Linear National Cost Equation

Table T-4.B-8: Means and Standard Errors of National Data

Table T-4.B-1a: DETERMINANTS OF AVERAGE OPERATING COSTS FOR ICF-SNF HOMES: THE COMPLETE COST MODEL; NATIONAL NCHS DATA

Explanatory Variable	Estimated Coefficient	Standard Error
Output: Scale		
CONSTANT	12.057	
AD1	-0.440	0.711
AD2	-0.167	0.151
AD3	-0.087xxx	0.021
AD4	-0.072xxx	0.025
AD5	-0.074xx	0.028
AD6	0.022x	0.028
OC1	0.043	0.010
OC2	0.157	0.074
OC3	-0.127	0.416
OC4	0.335	0.413
OC5	-0.412	0.495
OC6	0.135	0.321
Output: Patient Mix		
AGE	-0.007	0.072
FEMALE	0.005	0.031
Diagnosis		
ARTERY	0.065xxx	0.023
MENTAL	-0.001	0.026
STROKE	0.004	0.032
ACCIDENT	0.071x	0.042
NUTRIT	-0.022	0.050
OLD	0.017	0.029
RESPIRAT	-0.003	0.062
DIGEST	-0.031	0.077
NEOPLASM	0.042	0.065
HEART	0.039	0.034
Mental Status		
DEPRES	0.008	0.015
NERVOUS	0.013	0.017
ABUSE	-0.004	0.022
SENILE	-0.007	0.015
SLEEP	-0.006	0.016
Activities of Daily Living		
GROOMPD	0.007	0.019
GROOMD	0.022	0.027
DRESSPD	-0.019	0.025
DRESSD	0.004	0.030
EATPD	0.003	0.024
EATD	-0.198x	0.115
WALKPD	-0.024	0.022
WALKD	-0.016	0.029
BEDPD	-0.005	0.017
BEDD	0.057	0.044
CONTINPD	0.005	0.035
CONTIND	0.034	0.029

Table T-4.B-1a: (Cont'd.) DETERMINANTS OF AVERAGE OPERATING COSTS FOR ICF HOMES: THE COMPLETE COST MODEL; NATIONAL NCHS DATA

Explanatory Variable	Estimated Coefficient	Standard Error
Output: Services		
OT	0.462	0.792
RT	0.871	0.707
REHAB	-0.238	1.004
EDUC	0.374x	0.627
PT	1.544	0.910
SPHT	-0.216	0.769
COUNSEL	0.178	0.779
NASAL	0.347	0.218
BLOOD	-0.010	0.011
ENEMA	-0.005	0.019
CATH	0.020	0.045
BATH	0.0006	0.012
BOWEL	-0.030	0.019
OXYGEN	-0.174	0.117
BAND	0.010x	0.036
PULSE	0.021	0.011
HYPO	-0.011	0.012
IRRIG	-0.005	0.047
INTRAV	0.005	0.052
SUMSER	0.033	0.057
PONEB	0.011	0.021
PTHREEB	-0.039	0.033
Output: Case Flow		
ADMISPD	303.611xxx	93.941
DEATH	1.654	2.26
Input Prices		
RETAILWG	0.366	0.796
NURSEWG	3.040	0.382
SMSA	-0.502	0.739
Regulatory Variables		
PCTPRI	-0.036xx	0.016
PM1	-1.034	1.160
PM2	-0.590	1.229
CON	-0.369	1.150
LSC	1.238	0.761
LIM	-0.293	1.321

Table T-4.B-1a: (Cont'd.) DETERMINANTS OF AVERAGE OPERATING COSTS FOR ICF HOMES: THE COMPLETE COST MODEL; NATIONAL NCHS DATA

Explanatory Variable	Estimated Coefficient	Standard Error
Regulatory Variables (continued)		
STAFFA	1.824	1.354
STAFFB	-1.001	4.193
MA	-0.162	1.232
UNIF	0.784	1.045
CB	0.360	0.896
OCCUPEN	2.369x	1.074
NORTHEAST	3.368x	1.908
CENTRAL	1.168	1.373
SOUTH	1.768	4.441
Facility Characteristics		
PROFIT	-2.269x	1.258
NONPROFIT	-0.062	1.348
CHANGE	0.896	1.829
CHAIN	-1.279x	0.582
PCONTRACT	0.009	0.096
R^2	.671	
Adjusted R^2	.528	
F-ratio	4.69	
N	239	
Standard Error	4.304	

x significant at .10 level
xx significant at .05 level
xxx significant at .01 level

Table T-4.B-1b: DETERMINANTS OF AVERAGE OPERATING COSTS FOR ICF HOMES: THE COMPLETE COST MODEL; NATIONAL NCHS DATA

Explanatory Variable	Estimated Coefficient	Standard Error
Output: Scale		
CONSTANT	27.433	
AD1	-1.295xxx	0.236
AD2	-0.005	0.049
AD3	-0.029x	0.016
AD4	0.026	0.031
AD5	-0.026	0.041
AD6	-0.0002	0.019
OC1	0.086	0.091
OC2	-0.146	0.398
OC3	0.028	0.445
OC4	-0.197	0.513
OC5	-0.095	0.265
OC6	0.229	0.246
Output: Patient Mix		
AGE	-0.054	0.065
FEMALE	0.011	0.016
Diagnosis		
ARTERY	0.004	0.017
MENTAL	-0.004x	0.020
STROKE	-0.017	0.025
ACCIDENT	-0.0004	0.037
NUTRIT	-0.010	0.035
OLD	-0.007	0.020
RESPIRAT	-0.015	0.052
DIGEST	-0.036	0.055
NEOPLASM	-0.005	0.060
HEART	-0.0004	0.028
Mental Status		
DEPRES	0.020x	0.011
NERVOUS	-0.021	0.014
ABUSE	0.010	0.017
SENILE	0.002	0.012
SLEEP	-0.005	0.013
Activities of Daily Living		
GROOMPD	0.0007	0.014
GROOMD	0.009	0.019
DRESSPD	-0.006	0.019
DRESSD	-0.020	0.026
EATPD	0.007	0.023
EATD	-0.127	0.118
WALKPD	0.009	0.020
WALKD	0.008	0.026
BEDPD	-0.0005	0.016
BEDD	-0.008	0.032
CONTINPD	-0.053	0.036
CONTIND	-0.007	0.020

Table T-4.B-1b: (Cont'd.) DETERMINANTS OF AVERAGE OPERATING COSTS FOR ICF HOMES: THE COMPLETE COST MODEL; NATIONAL NCHS DATA

Explanatory Variable	Estimated Coefficient	Standard Error
Output: Services		
OT	0.260	0.839
RT	-0.666	0.685
REHAB	-1.369	1.373
EDUC	1.065x	0.557
PT	1.561x	0.618
SPHT	-2.131xx	0.899
COUNSEL	0.575	0.530
NASAL	0.400	0.299
BLOOD	0.011	0.008
ENEMA	0.056xx	0.022
CATH	0.062	0.065
BATH	-0.005	0.009
BOWEL	-0.006	0.013
OXYGEN	0.057	0.157
BAND	0.031	0.042
PULSE	-0.008	0.008
HYPO	0.004	0.012
IRRIG	0.007	0.052
INTRAV	-0.001	0.092
SUMSER	0.015	0.059
PONEB	0.044x	0.016
PTHREEB	-0.014	0.032
Output: Case Flow		
ADMISPD	531.366xxx	111.752
DEATH	4.012xx	1.778
Input Prices		
RETAILWG	-0.030	0.596
NURSEWG	1.342xxx	0.352
SMSA	-0.392	0.534
Regulatory Variables		
PCTPRI	0.005	0.011
PM1	-1.380	1.129
PM2	0.129	1.137
CON	1.316	1.109
LSC	0.743	0.670
LIM	-2.208	1.682

Table T-4.B-1b: (Cont'd.) DETERMINANTS OF AVERAGE OPERATING COSTS FOR ICF HOMES: THE COMPLETE COST MODEL; NATIONAL NCHS DATA

Explanatory Variable	Estimated Coefficient	Standard Error
Regulatory Variables (continued)		
STAFFA	1.113	1.044
STAFFB	0.246	4.679
MA	0.270	1.075
UNIF	0.217	1.044
CB	0.477	0.849
OCCUPEN	1.360	1.011
NORTHEAST	0.343	2.159
CENTRAL	1.646	1.353
SOUTH	1.649	1.976
Facility Characteristics		
PROFIT	0.445	1.026
NONPROFIT	1.797x	1.014
CHANGE	-2.137	3.170
CHAIN	-1.074x	0.552
PCONTRACT	0.157x	0.083

R^2	.640
Adjusted R^2	.481
F-ratio	4.012
N	296
Standard Error	3.368

x significant at .10 level
xx significant at .05 level
xxx significant at .01 level

Table T-4.B-1c: DETERMINANTS OF AVERAGE OPERATING COSTS FOR SNF HOMES: THE COMPLETE COST MODEL; NATIONAL NCHS DATA

Explanatory Variable	Estimated Coefficient	Standard Error
Output: Scale		
CONSTANT	27.66	.499
AD1	-.682	.085
AD2	-.239xx	.016
AD3	-.025	.022
AD4	-.023	.030
AD5	008	.011
AD6	-.002	.055
OC1	-.086	.338
OC2	-.172	.360
OC3	.058	.406
OC4	-.174	.247
OC5	.015	.289
OC6	-.242	
Output: Patient Mix		
AGE	.028	.057
FEMALE	-.030	.021
Diagnosis		
ARTERY	.036x	.019
MENTAL	.033	.027
STROKE	.073xxx	.024
ACCIDENT	.014	.033
NUTRIT	-.020	.039
OLD	.004	.021
RESPIRAT	.070x	.052
DIGEST	.112	.051
NEOPLASM	.006	.049
HEART	.012	.031
Mental Status		
DEPRES	.001	.013
NERVOUS	.016	.013
ABUSE	-.032x	.019
SENILE	-.008	.013
SLEEP	-.003	.015
Activities of Daily Living		
GROOMPD	.016	.016
GROOMD	-.033	.021
DRESSPD	.019	.021
DRESSD	-.001	.026
EATPD	.037	.021
EATD	.109*	.071
WALKPD	.017	.019
WALKD	.019	.022
BEDPD	-.010	.014
BEDD	.034	.036
CONTINPD	.045	.027
CONTIND	.025	.017

Table T-4.B-1c: (Cont'd.) DETERMINANTS OF AVERAGE OPERATING COSTS FOR SNF HOMES: THE COMPLETE COST MODEL; NATIONAL NCHS DATA

Explanatory Variable	Estimated Coefficient	Standard Error
Output: Services		
OT	1.232x	.590
RT	.292	.571
REHAB	.656	.956
EDUC	.352x	.538
PT	1.853	.883
SPHT	.592	.637
COUNSEL	.835	.673
NASAL	-.034	.070
BLOOD	.007	.008
ENEMA	.013	.015
CATH	.012	.039
BATH	.012	.009
BOWEL	.009	.013
OXYGEN	.167	.120
BAND	.018	.024
PULSE	.001	.008
HYPO	.015	.012
IRRIG	.0023	.029
INTRAV	.013	.087
SUMSER	.070x	.039
PONEB	.015	.019
PTHREEB	.032	.029
Output: Case Flow		
ADMISPD	403.246xxx	62.453
DEATH	-.095	1.186
Input Prices		
RETAILWG	.913	.579
NURSEWG	3.239xxx	.326
SMSA	-.541	.646
Regulatory Variables		
PCTPRI	.011	.011
PM1	-.517	.967
PM2	-1.209	.963
CON	-1.570	1.299
LSC	.013	.290
LIM	-1.810	1.154
STAFFA	3.207x	1.540
STAFFB	2.222	2.946
MA	.207x	1.238
UNIF	2.141	1.032
CB	-1.032	.850
OCCUPEN	-.698x	1.140
NORTHEAST	3.356	1.811
CENTRAL	-.964	1.623
SOUTH	.683	1.450

Table T-4.B-1c: (Cont'd.) DETERMINANTS OF AVERAGE OPERATING COSTS FOR SNF HOMES: THE COMPLETE COST MODEL; NATIONAL NCHS DATA

Explanatory Variable	Estimated Coefficient	Standard Error
Facility Characteristics		
PROFIT	-4.308xxx	1.162
NONPROFIT	-2.955xx	1.243
CHANGE	.673	2.150
CHAIN	.247	.526
PCONTRACT	.001	.050

R^2	.616
Adjusted R^2	.543
F-ratio	8.4
N	549
Standard error	5.314

x significant at .10 level
xx significant at .05 level
xxx significant at .01 level

Table T-4.B-2

NURSING STAFF HOURS PER PATIENT DAY NEEDED TO CARE FOR SELECTED PATIENT TYPES*

Aide or Orderly Hours	Patient Type
.05	Needs help in getting and/or putting away equipment or needs help with bathing but can wash face and hands, brush teeth, comb hair, and shave or apply makeup
.25	Percentage of patients who need partial assistance in dressing
.5	Percentage of patients who stay partly or completely undressed
.1	Feeds self with minor assistance
.35	Feeds self with major assistance
.6	Receives either intravenous feeding or tube feeding
.35	Walks with assistance of one or two persons
.45	Up in a chair with assistance once or twice in an eight hour period
.5	Is bedfast
.1	Is in control of bladder only or bowels only
.1	Is not in control of bowel and/or bladder retraining
.05	Percentage of patients with mental disorders (e.g., mental retardation, psychoses, neuroses, mental illness, emotional problems
.1	Percentage of patients with diseases of the nervous system and sense organs
.1	Receives bowel and/or bladder retraining

*The aide or orderly hours needed are taken from Winn, 1975.

Table T-4.B-3

NURSING HOMES PER PATIENT DAY NEEDED TO CARE FOR PATIENTS DEPENDENT ON ACTIVITIES OF DAILY LIVING*

RN, LPN, or Aide Hours	Patient Type
.29	Needs help in getting and/or putting away equipment or needs help with bathing can wash face and hands, brush teeth, comb hair, and shave or apply makeup
.36	Needs help in at least two of the following activities: washing face and hands, brushing teeth, or dentures, combing hair, and shaving or applying makeup
.83	Needs partial assistance in dressing
.75	Feeds self with major assistance
.941	Receives either intravenous feeding or tube feeding
1.0	Receives nurses' assistance in going to the toilet room (do not count on use of cane, walker, or other object of support), using bedpan or commode at night, or cleaning self, or arranging clothes after elimination
.91	Unable to go to the toilet room for the elimination process
.4995	Walks with assistance of one or two persons
.4	Up in a chair with assistance once or twice in an eight hour period
.2	Is bedfast
.166	Is in control of bladder only or bowels only
.983	Is not in control of bowel and/or bladder retraining

*The nursing hours needed were determined by June Jensen, R.N., on the basis of her expert opinion from working in nursing homes. These numbers are needed to test the potential utility of this technique rather than definitive statements.

Table T-4.B-4: Residual Analysis of Complete Model: Bias Analysis of Selected Nursing Home Characteristics by Size of Standardized Residual

Standardized Residual Range	Frequency	Average Operating Expenses		Bedsize	
		Mean	St. Dev.	Mean	St. Dev.
lowest thru -5	1	12.25	0.0	104.00	0.0
-5.01 thru -4	6	17.36	11.68	114.16	70.63
-4.01 thru -3	2	10.42	.34	43.00	9.89
-3.01 thru -2	6	13.80	4.91	150.50	157.82
-2.01 thru -1	96	14.74	7.81	132.68	132.77
-1.01 thru 0	507	14.74	9.12	128.04	116.06
0.01 thru 1	417	13.90	7.24	122.32	98.60
1.01 thru 2	59	13.41	4.92	134.35	121.79
2.01 thru 3	17	18.99	17.32	149.64	107.05
3.01 thru 4	8	22.86	25.54	109.16	60.89
4.01 thru 5	2	24.71	16.99	90.00	14.14
5.01 thru highest	6	20.36	6.09	87.00	49.50

Standardized Residual Range	Frequency	Census		Occupancy Rate	
		Mean	St. Dev.	Mean	St. Dev.
lowest thru -5	1	98.00	0.0	94.00	0.0
-5.01 thru -4	6	108.16	71.38	96.66	38.68
-4.01 thru -3	2	40.50	13.43	93.00	98.00
-3.01 thru -2	6	145.16	154.37	95.50	3.10
-2.01 thru -1	96	120.04	104.78	93.14	100.54
-1.01 thru 0	507	119.35	107.61	93.28	117.59
0.01 thru 1	417	118.88	88.63	92.39	111.74
1.01 thru 2	59	124.93	117.80	92.06	90.30
2.01 thru 3	17	141.11	97.77	95.76	44.06
3.01 thru 4	8	99.50	51.51	93.12	40.41
4.01 thru 5	2	84.50	16.26	93.50	12.50
5.01 thru highest	6	80.83	49.62	90.33	163.06

Table T-4.B-5: Residual Analysis of Complete Model: Heteroskedasticity Distribution of Standardized Residuals by Selected Facility Characteristics

	Residuals		
	Mean	St. Dev.	N
SMSA			
Part of an SMSA	.04	1.04	729
Not part of an SMSA	-.08	.90	398
OWNERSHIP			
State	-.14	.57	10
County	.22	.80	80
City	.10	.59	16
City-County	.79	.09	4
Hospital District	-.55	.44	4
Church Related	-.06	.80	80
Nonprofit Corporation	.06	.06	128
Other Nonprofit Ownership	.01	.92	8
Individual	-.05	.93	106
Partnership	.18	1.46	79
Corporation	.01	.97	612
CERTIFICATION			
Medicare Only	-.06	1.48	37
SNY Only	-.17	.86	194
ICF Only	.02	.81	296
Medicare and SNH	.06	.19	361
Medicare and ICF	.30	.77	3
SNF and ICF	.01	.75	136
Medicare, SNF and ICF	.03	1.01	100
REGION			
Northeast	.01	1.29	269
Central	.02	.84	365
South	.00	.82	301
West	-.05	1.06	192
BEDSIZE			
Less than 50	-.02	1.08	152
51 thru 99	-.04	.92	390
100 thru 199	.07	1.10	414
200 or greater	-.05	.80	171

Table T-4.B-6: FREQUENCY DISTRIBUTIONS OF SELECTED DISCRETE VARIABLES WITH THE STANDARDIZED RESIDUALS

	N	Lowest thru-5	-5.01 thru-4	-4.01 thru-3	-3.01 thru-2	-2.01 thru-1	-1.01 thru-0	0.01 thru-1	1.01 thru-2	2.01 thru-3	3.01 thru-4	4.01 thru	thru highest
SMSA													
Part of an SMSA	729	0	5	0	4	63	327	258	48	15	3	1	5
Not part of an SMSA	398	1	1	2	2	33	180	159	11	2	5	1	1
OWNER													
Profit	797	1	5	2	4	65	350	301	45	12	6	1	5
Non-profit	216	0	1	0	1	18	98	80	12	3	1	1	1
Government	114	0	0	0	1	13	59	36	2	2	1	0	0
CERTIFICATION													
SNF	592	1	4	1	4	60	269	204	28	10	5	1	5
ICF-SNF	239	0	1	0	1	16	115	85	14	5	0	1	1
ICF	296	0	1	1	1	20	123	128	17	2	3	0	0
REGION													
Northeast	269	0	2	1	4	32	116	84	17	8	2	1	2
Central	365	0	1	1	0	21	176	139	16	6	3	1	1
South	301	1	2	0	0	25	122	129	16	3	3	0	0
West	192	0	1	0	2	18	93	65	10	0	0	0	3
BEDSIZE													
Less than 50	152	0	1	1	0	13	72	55	6	1	1	0	2
50 thru 99	390	0	2	1	3	37	175	143	19	4	4	1	1
100 thru 199	414	1	2	0	2	32	176	161	26	8	2	1	3
200 or more	171	0	1	0	1	14	84	58	8	4	1	0	0

Table T-4.B-7: DETERMINANTS OF AVERAGE OPERATING COSTS: LOG-LINEAR MODEL; NATIONAL NCHS DATA

Explanatory Variable	Regression Coefficient	Standard Error
Output: Scale		
CONSTANT	2.4603	
AD1	-0.0413xxx	0.0118
AD2	-0.0037	0.0024
AD3	-0.0019xxx	0.0005
AD4	0.0006	0.0008
AD5	-0.0012	0.0010
AD6	0.0005	0.0004
OC1	0.0001	0.0021
OC2	-0.0078	0.0114
OC3	0.0094	0.0122
OC4	-0.0200	0.0143
OC5	0.0030	0.0083
OC6	-0.0011	0.0093
Output: Patient Mix		
AGE	-0.0007	0.0020
FEMALE	0.0004	0.0007
Diagnosis		
ARTERY	0.0023xxx	0.0006
MENTAL	-0.0254	0.0008
STROKE	0.0015x	0.0008
ACCIDENT	0.0020x	0.0012
OLD	0.0004	0.0007
RESPIRAT	0.0008	0.0017
DIGEST	0.0008	0.0018
NEOPLASM	0.0004	0.0017
HEART	0.0358	0.0010
Mental Status		
DEPRES	0.0011xxx	0.0004
NERVOUS	-0.0001	0.0005
ABUSE	-0.0002	0.0006
SENILE	0.0001	0.0004
SLEEP	-0.0009x	0.0005
Activities of Daily Living		
GROOMPD	0.0005	0.0005
GROOMD	0.0000	0.0007
DRESSPD	0.0000	0.0007
DRESSD	-0.0006	0.0009
EATPD	0.0010	0.0007
EATD	0.0005	0.0028
WALKPD	0.0010	0.0006
WALKD	0.0015x	0.0008
BEDPD	0.0003	0.0005
BEDD	0.0009	0.0012
CONTINPD	-0.0016	0.0010
CONTIND	-0.0005	0.0006

Table T-4.B-7: DETERMINANTS OF AVERAGE OPERATING COSTS: LOG-LINEAR MODEL; NATIONAL NCHS DATA (Cont'd.)

Explanatory Variable	Regression Coefficient	Standard Error
Output: Services		
OT	0.0595xxx	0.0219
RT	0.0399xx	0.0199
REHAB	-0.0216	0.0323
EDUC	0.0328x	0.0178
PT	0.0962xxx	0.0250
SPHT	0.0120	0.0224
COUNSEL	0.0231	0.0209
NASAL	-0.0010	0.0030
BLOOD	0.0006xx	0.0003
ENEMA	0.0009x	0.0005
CATH	0.0015	0.0014
BATH	0.0001	0.0003
BOWEL	0.0001	0.0005
OXYGEN	0.0009	0.0039
BAND	0.0010	0.0009
PULSE	0.0003	0.0003
HYPO	-0.0001	0.0004
IRRIG	0.0004	0.0011
INTRAV	-0.1546	0.2181
SUMSER	0.0032xx	0.0015
PONEB	0.0010x	0.0006
PTHREEB	0.0008	0.0010
Output: Case Flow		
ADMISPD	21.6305xxx	2.4129
DEATH	0.0393	0.0464
Input Prices		
RETAILWG	0.0556xxx	0.0202
NURSEWG	0.1486xxx	0.0108
SMSA	-0.0598xxx	0.0201
Regulatory Variables		
ICF	-0.1728xxx	0.0267
ICF-SNF	-0.0252	0.0218
PCTPRI	0.0004	0.0004
PTCARE	0.0011	0.0010
PM1	-0.0761xxx	0.0293
PM2	-0.0503	0.0316
CON	0.0298	0.0331
LSC	0.0505xx	0.0220
LIM	0.0120	0.0360

Table T-4.B-7: DETERMINANTS OF AVERAGE OPERATING COSTS: LOG-LINEAR MODEL; (Cont'd.) NATIONAL NCHS DATA

Explanatory Variable	Regression Coefficient	Standard Error
Regulatory Variables (continued)		
STAFFA	0.1034xxx	0.0759
STAFFB	-0.0528	0.1010
MA	0.0202	0.0336
UNIF	0.0444	0.0294
CB	0.0233	0.0244
OCCUPEN	0.1082xxx	0.0302
NORTHEAST	0.2609xxx	0.0514
CENTRAL	0.1361xxx	0.0431
SOUTH	0.1789xxx	0.0428
Facility Characteristics		
PROFIT	-0.1354xxx	0.0363
NONPROFIT	-0.0086	0.0383
CHANGE	0.0443	1.0638
CHAIN	-0.0394xx	0.0172
PCONTRACT	0.0011	0.0020

Associated with predicted antilogarithms:	
R^2	0.452
Associated with predicted logarithms:	
R^2	0.647
Adjusted R^2	0.616
F-ratio	20.871
N	1127
Standard Error	0.264

x significant at 0.10 level
xx significant at 0.05 level
xxx significant at 0.01 level

Table T-4.B-8: MEANS AND STANDARD ERRORS OF NATIONAL DATA
(based on 1127 cases)

Variable	Mean	Standard Dev.
Output: Scale		
AD	114.786	2.503
Output: Utilization		
OCC	0.924	0.003
Output: Patent Mix		
AGE	79.5432	4.8303
FEMALE	71.0025	13.6758
DIAGNOSIS		
ARTERY	24.0209	18.9500
MENTAL	9.3767	14.8149
STROKE	11.3329	12.2222
ACCIDENT	4.7918	8.2051
NUTRIT	4.1105	6.8909
OLD	12.6520	15.7207
RESPIRAT	2.0742	4.9984
DIGEST	1.9224	4.6941
NEOPLASM	2.3766	5.1784
HEART	5.4374	9.4919
MENTAL STATUS		
DEPRES	40.3160	24.0962
NERVOUS	43.3495	23.7424
ABUSE	16.4053	15.3106
SENILE	58.7197	24.4199
SLEEP	19.0057	20.7278
ACTIVITIES OF DAILY LIVING		
GROOMPD	28.3114	20.6191
GROOMD	45.6657	24.1200
DESSPD	44.9594	23.7974
DRESSD	17.1405	19.8107
EATPD	18.0828	16.1133
EATD	0.8053	3.0783
WALKPO	32.7445	21.7582
WALKD	22.7650	20.0569
BEDPD	26.3705	20.4521
BEDD	5.0188	9.0212
CONTINPD	5.8364	8.8878
CONTIND	29.6203	19.6768

Table T-4.B-8: MEANS AND STANDARD ERRORS OF NATIONAL DATA
(Cont'd.) (based on 1127 cases)

Variable	Mean	Standard Dev.
OUTPUT: SERVICES		
OT	0.3496	0.4771
RT	0.4161	0.4931
REHAB	0.0710	0.2569
EDUC	0.4153	0.4930
PT	0.7063	0.4557
SPHT	0.4321	0.4956
COUNSEL	0.6832	0.4654
NASAL	0.7560	3.0688
BLOOD	55.3274	37.7352
ENEMA	14.0000	16.6005
CATH	5.7799	8.1752
BATH	26.4303	29.4704
BOWEL	14.1792	18.3109
OXYGEN	1.3949	2.4778
BAND	9.0745	10.9587
PULSE	50.4490	39.6881
HYPO	19.5972	21.9205
IRRIG	8.5235	10.0936
INTRAV	0.0076	0.0405
SUMSER	4.1460	6.7532
PONEB	11.6189	18.5533
PTHREEB	8.8000	9.6598
OUTPUT: CASE FLOW		
ADMISPD	0.0019	0.0047
DEATH	0.3080	0.2329
INPUT PRICES		
RETAILWG	2.6213	0.4214
NURSEWG	2.3824	0.9206
SMSA	0.3283	0.4698
REGULATORY VARIABLES		
ICF	0.254	0.435
ICF-SNF	0.258	0.434
PCTPRI	38.3203	27.7014
PM1	0.3931	0.4887
PM2	0.3256	0.4688
CON	0.4667	0.4991
LSC	0.3828	0.4812
LIM	0.8438	0.3632
STAFFA	0.1136	0.3138
STAFFB	0.1080	0.1126
MA	0.3265	0.4692
UNIF	0.3159	0.4651
CB	0.4880	0.5001

Table T-4.B -8: MEANS AND STANDARD ERRORS OF NATIONAL DATA
(Cont'd.) (based on 1127 cases)

Variable	Mean	Standard Dev.
OCCUPEN	0.4738	0.4995
NORTHEAST	0.2059	0.4045
CENTRAL	0.2910	0.4544
SOUTH	0.2884	0.4532
FACILITY CHARACTERISTICS		
PROFIT	0.7959	0.4032
NONPROFIT	0.1278	0.3340
CHANGE	0.0310	0.1349
CHAIN	0.5803	0.4937
PCONTRACT	2.6762	4.2879
COST COMPONENTS		
PAYEXP	10.423	0.165
NUREXP	5.853	0.085
FOOD	1.423	0.023
MAINT	0.271	0.010
LAUN	0.179	0.008

APPENDIX T-4.C

COST COMPONENT ANALYSIS

T-4.C.1 COST COMPONENT ANALYSIS

Previous sections have analyzed the determinants of nursing home operating costs. This section considers the subsidiary issue of the determinants of the major sub-components which make up operating costs. In particular, it discusses linear regression results obtained by modeling payroll expenses, nursing expenses food and dietary expenses laundry and linen expenses and expenses for the maintenance of the building, grounds, and equipment, all on a per patient day basis.[1] The R^2 for these cost component functions ranged from a high of 0.64 for the nursing expenses category to a low of 0.26 for both the maintenance category and the laundry and linen category. The regression results are reported in Table T-4.B-1. These explanatory variables for each of the cost component models other than patient days and the occupancy rate were entered only if their F value was greater than 1. Exploratory regression analyses suggested that the gain in accuracy by using the complete specification was trivial. Since the results here are for illustration only, and (in the context of a formula prospective reimbursement mechanism) do not enter directly into a reimbusement formula, the discussion here considers only the policy specification approach.

The results obtained in the five departmental analyses are strikingly similar to the larger cost analysis in two ways. First, the signs of the coefficients of the explanatory variables are basically the same as those for the operating cost analyses and second, the set of variables entered across the different departmental analyses is about the same as the set entered in the policy model of operating expenses.

Payroll expenses are defined as wages and salaries (gross amount including employees' FICA, vacation and sick pay, taxes, etc.) and fringe benefits (state unemployment, group health insurance and all other payroll

---

[1] Payroll expenses account for 72.7 percent of operating costs; nursing expenses account for 56.1 percent of payroll expenses. Food and dietary expenses account for 9.9 percent of operating expenses. Laundry and linen expenses account for 1.3 percent of operating expenses. And maintenance expenses accounts for 1.9 percent of operating expenses. Together, payroll, food, laundry and maintenance account for 85.8 percent of operating costs.

and non-payroll benefits for all employees on the staff). Nursing expenses measure the wages and salaries of RNs, LPNs, practical nurses, aides, orderlies, student nurses, and other nursing staff. Food and dietary expenses are defined as the costs of dietary and food items purchased for preparation on the home's premise. Contracted services are omitted. Laundry and linen expenses include the cost of only non-contracted services. Maintenance expenses include the costs of plumbers, electricians, ground maintenance, carpenters, general repairmen, and similar non-contracted maintenance services.

T-4.C.2 Regression Results

The individual cost component analyses are discussed first. Then the differential impacts of the types of causal variables across the alternative types of cost centers are considered. The four basic groups of variables in the total average operating cost model (output related, input price related, regulatory related, and those variables which further characterize facilities) enter these cost-center equations. However, some modifications are made to reflect the appropriateness of the explanatory variables to the particular cost center; for example, in the food equation, the input price variable, kitchen help wages (COOKWG) is included and the variable NURSEWG is deleted.

T-4.C.2.1 Payroll Expenses

Turning to payroll expenses (Table T-4.C-1), the R^2 for this equation is .599. The main output variable is average daily census. With regard to patient days there are, as found earlier, slight economies of scale. No significant consistent relationship is evident between occupancy rates and payroll expenses. The admission rate (i.e., the number of admissions per patient day) was allowed to enter the payroll expenses function to measure the cost of an admission in this category. A strong statistical relationship exists between this variable and payroll expenses. Examining the coefficients of ADMISPD in the total average operating cost equation and in the payroll equation reveals that the total cost of an admission (approximately $415) is composed primarily of payroll expenses (about $297).

Case mix and service intensity do not appear as strong cost determinants as indicated by both the small magnitude and marginal significance of the corresponding coefficients. Very few measures of service intensity entered the model.

Table T-4.C-1: DETERMINANTS OF PER DIEM PAYROLL EXPENSES: NATIONAL NCHS DATA

Explanatory Variable	Estimated Coefficient	Standard Error
Output: Scale		
CONSTANT	15.322	
AD1	-0.656xxx	0.146
AD2	-0.057x	0.030
AD3	-0.023xxx	0.006
AD4	0.013	0.009
AD5	0.015	0.012
AD6	0.011xx	0.004
OC1	0.002	0.025
OC2	-0.089	0.140
OC3	0.109	0.150
OC4	-0.184	0.177
OC5	0.027	0.103
OC6	0.038	0.115
Diagnosis		
ARTERY	-0.016xx	0.006
MENTAL	-0.020xx	0.007
STROKE	0.010	0.009
ACCIDENT	0.018	0.013
NUTRIT	-0.020	0.015
NERVES	-0.014	0.012
Mental Status		
DEPRES	0.009x	0.004
Activities of Daily Living		
GROOMPD	0.007	0.005
DRESSD	-0.009	0.006
EATPD	0.013x	0.007
EATD	0.050	0.034
BEDPD	0.022x	0.013
CONTINPD	0.017	0.011
Output: Services		
OT	0.576	0.264
RT	0.372	0.242
EDUC	0.221	0.217
PT	0.742xx	0.293
BLOOD	0.008xx	0.002
ENEMA	0.012xx	0.006
INTRAV	-0.032	0.002
SUMSER	0.053xxx	0.017

Table T-4.C-1: DETERMINANTS OF PER DIEM PAYROLL EXPENSES:
(Cont'd.) NATIONAL NCHS DATA

Explanatory Variable	Estimated Coefficient	Standard Error
Output: Case Flow		
ADMISPD	297.313xxx	27.170
Input Prices		
RETAILWG	0.510x	0.248
NURSEWG	2.213xxx	0.129
Regulatory Variables		
ICF	-1.802xxx	0.310
ICF-SNF	-0.368	0.259
PM1	0.024xxx	0.012
PM2	-1.383xxx	0.330
CON	-1.363xxx	0.340
STAFFA	0.634	0.420
MA	0.732x	0.309
UNIF	0.791xx	0.336
OCCUPEN	0.674x	0.320
NORTHEAST	3.305xxx	0.420
CENTRAL	1.582xxx	0.468
SOUTH	1.995xxx	0.430
Facility Characteristics		
PROFIT	1.480xxx	0.399
NONPROFIT	0.483	0.468
CHANGE	0.871	0.788
PCONTRACT	0.042x	0.024

R^2	.599
Adjusted R^2	.580
F-ratio	30.289
N	1127
Standard Error	3.339

x significant at .10 level
xx significant at .05 level
xxx significant at .01 level

In contrast, input prices appear as important cost determinants. Both the average nurse's wage, NURSEWG, and the average county wage, WAGE, have a significant positive relationship to payroll expenses.

With respect to regulatory related variables, the certification dummies, both ICF and ICFSNF, entered the model. ICF certified homes have significantly lower payroll costs per patient day than SNF homes--about $1.80 less. ICF-SNF homes are lower by $0.368. All regional dummy variables enter the model as significant. Ordering regions of the country in terms of most costly to least costly yields the same order as in the overall average operating cost equation: Northeast, South, Northcentral, and West.

With regard to state regulations, the requirement of maintaining a uniform chart of accounts adds about $0.79 to payroll expenses per patient day (indicated by the coefficient of UNIF) while the requirement that an RN or LPN be on duty daily irrespective of the size of the home adds about $0.63 to payroll expenses per patient day (indicated by the coefficient of STAFFA). The coefficients of the payment mechanism dummy variables, PM1 and PM2, suggest that prospective payment systems contribute about $1.38 more to payroll per diem costs than do cost related systems (either reasonable, actual, or retrospective), and about $1.36 more to payroll per diem expenses over flat rate systems. Pre-admission patient assessment relative to post-admission assessment adds about $0.73 to payroll expenses (indicated by the coefficient of MA). The results also suggest that occupancy floors add approximately $0.67 to payroll expenses per patient day (indicated by the coefficient of OCCPEN).

Finally, with regard to ownership ceteris paribus, government homes have the highest per diem payroll expenses. Other non-profit homes have per diem payroll expenses about $0.48 less per patient day than government homes; for-profit homes have the lowest per diem payroll expenses--about $1.48 less than government homes and $1.00 less than non-profit homes.

T-4.C.2.2 Nursing Expenses

For the components of nursing expenses per patient day (Table T-4.C-2), the regression results are similar to those discussed for the payroll expenses equation. This is expected since nursing expenses account for the majority (56 percent) of payroll expenses. The nursing cost equation has an R^2 of 0.643. There are few noteworthy differences from the payroll regression.

Table T-4.C-2: DETERMINANTS OF PER DIEM NURSING COSTS: NATIONAL NCHS DATA

Explanatory Variable	Estimated Coefficient	Standard Error
Output: Scale		
CONSTANT	9.012	
AD1	-.420xxx	.073
AD2	-.058xxx	.015
AD3	-.010xxx	.003
AD4	.0006	.004
AD5	.008	.006
AD6	-.006xxx	.002
OC1	-.004	.012
OC2	.0009	.070
OC3	.065	.075
OC4	-.129	.088
OC5	.048	.052
OC6	-.009	.057
Output: Patient Mix		
FEMALE	.006	.004
Diagnosis		
ARTERY	.008xxx	.003
MENTAL	.009xx	.004
STROKE	.009xxx	.004
ACCIDENT	.010	.006
NEOPLASM	.014	.010
OTHARTERY	.008	.007
Mental Status		
NERVOUS		
ABUSE	.006xx	.003
Activities of Daily Living		
GROOMPD	.004	.002
DRESSPD	.006xxx	.002
DRESSD	.003	.003
EATPD	.005	.003
EATD	.018	.017
BEDPD	.003	.002
CONTIND	.009	.005
Output: Services		
OT	.155	.132
RT	.185	.122
REHAB	-.311	.201
PT	.428xxx	.146
NASAL	.032x	.018
BLOOD	.002	.001
ENEMA	.006xx	.003
OXYGEN	.040x	.023
PULSE	.001	.001
IRRIG	-.009	.006
INTRAV	.002	.001
SUMSER	.014	.008
PTHREED	.012x	.003

Table T-4.C-2: DETERMINANTS OF PER DIEM NURSING COSTS: NATIONAL NCHS DATA (Cont'd.)

Explanatory Variable	Estimated Coefficient	Standard Error
Output: Case Flow		
ADMISPD	142.339xxx	13.30
Input Prices		
RETAILWG	.319xx	.125
NURSEWG	1.495xxx	.066
SMSA		
Regulatory Variables		
ICF	-1.093xxx	.155
ICF-SNF	-.254xx	.130
PM1	-.667xxx	.166
PM2	-.764xxx	.176
LSC	.209	.127
STAFFA	.131	.216
MA	.206	.156
UNIF	.239	.175
OCCUPEN	.291x	.171
NORTHEAST	1.518xxx	.216
CENTRAL	.780xxx	.235
SOUTH	1.118xxx	.216
Facility Characteristics		
PROFIT	-.586xx	.221
NONPROFIT	-.377	.236
R^2	.643	
Adjusted R^2	.623	
F-ratio	33.137	
N	1127	
Standard Error	1.672	

x significant at .10 level
xx significant at .05 level
xxx significant at .01 level

Note the slight but significant economies of scale with regard to patient days provided. The results suggest that of the $297 payroll cost of an admission, about 40 percent (or $142) is nursing expenses.

With regard to other variables in the model, the signs of the coefficients and the pattern of significance of the payroll and nursing cost models are quite similar; only the magnitude of the coefficients differ. In general, the coefficients in the nursing cost model are smaller in absolute value.

T-4.C.2.3 Food and Dietary Expenses

The regression results for food and dietary expenses (Table T-4.C-3) are somewhat different than those obtained for both the payroll and nursing expense models. The equation accounts for 33.8 percent of the variation in per diem food and dietary expenses.

With regard to output, there are slight but significant economies of scale for an average daily census less than 150. As with the payroll and nursing cost models, no definite pattern occurs with respect to the occupancy rate. It is interesting that patient diagnoses appear more significantly related to food and dietary costs (perhaps because of the costs of special food preparation) than does the level of patient debility in terms of activities of daily living. Measures of service intensity were not allowed to enter this cost-component model.

As with the cost-component functions already examined, input price is the most statistically significant variable. For this equation, a kitchen wage proxy variable was constructed as the average hourly rate of cooks and dieticians in the home.

With regard to state regulatory variables, only the presence of life safety codes appear as significant, adding about $0.08 per patient day to food and dietary expenses. Again, the certification variables are significant indicating that ICF homes have the lowest per diem food and dietary expenses, mixed ICF-SNF homes the next lowest, and SNF homes the highest food and dietary costs. The Northeast and the South have significantly higher food and dietary expenses when compared to the West. Finally, volunteer and government homes have about $0.22 higher food and dietary costs per patient day when compared with for-profit homes.

Table T-4.C-3: DETERMINANTS OF PER DIEM FOOD AND DIETARY COSTS; NATIONAL NCHS DATA

Explanatory Variable	Estimated Coefficient	Standard Error
Output: Scale		
CONSTANT	4.410	
AD1	-.133xxx	.025
AD2	-.001	.005
AD3	.003xxx	.001
AD4	.001	.001
AD5	.001	.002
AD6	.002	.0007
OC1	-.013xx	.004
OC2	.025	.024
OC3	-.027	.026
OC4	-.037	.030
OC5	.049xxx	.017
OC6	.036x	.019
Output: Patient Mix		
FEMALE	.001	.001
Diagnosis		
ARTERY	.002xx	.001
MENTAL	-.019	.001
STROKE	.005xxx	.001
ACCIDENT	.003	.002
DIGEST	.009	.003
OTHARTERY	.003	.002
Activities of Daily Living		
GROOMPD	.002xx	.0008
DRESSPD	-.001	.001
EATD	.007	.005
BEDPD	.003	.002
CONTINPD	-.003	.002
Output: Services		
PONEB	.004xxx	.001
PTHREEB	.002	.002
Input Prices		
RETAILWG	.132xxx	.043
COOKWG	.165xxx	.011
SMSA	-.049	.041

Table T-4.C-3: DETERMINANTS OF PER DIEM FOOD AND DIETARY COSTS; (Cont'd.) NATIONAL NCHS DATA

Explanatory Variable	Estimated Coefficient	Standard Error
Regulatory Variables		
ICF	-.175xxx	.050
ICF-SNF	.075x	.045
PCTPRI	.001x	.0007
PCTCARE	.006xxx	.001
PM1	.077	.049
PM2	.095	.059
CON	.080	.042
LIM	-.039	.057
OCCUPEN	.084	.053
NORTHEAST	.270xxx	.063
Facility Characteristics		
PROFIT	-.221xxx	.052
R^2	.338	
Adjusted R^2	.312	
F-ratio	13.157	
N	1127	
Standard Error	.580	

x significant at .10 level
xx significant at .05 level
xxx significant at .01 level

T-4.C.2.4 Laundry Expenses

The estimated model for laundry and linen per diem expenses (Table T-4.C-4) accounted for 26.7 percent of the cost variation. Neither economies of scale nor utilization are evident. The admissions rate is significant. With regard to case mix, the percentage of female patients, average patient age, and a few ADL variables appear as significant. As with the food and dietary component model, measures of service intensity were not allowed to enter the laundry and linen equation.

A proxy housekeeping and maintenance personnel wage variable, HOUSEWG, was allowed to enter the model. Analogous to all earlier results it is the most statistically significant explanatory variable. However, the average county retail wage variable is insignificant.

As before, the coefficients of the certification dummy variables, ICF and ICFSNF, are significant suggesting the same relationship that has appeared earlier. ICF homes are less expensive than mixed ICF-SNF homes, which in turn are less expensive than SNF homes. There are no other significant regulatory related variables in the model. Note that ownership type did not enter the model.

T-4.C.2.5 Maintenance Expenses

The final cost component examined was per diem maintenance expenses (Table T-4.C-5). The equation accounts for 26.3 percent of the cost variation. The regression results are extremely similar to the laundry and linen equation results. Briefly, economies of scale and for utilization are not evident. The case mix variables which entered are, for the most part, insignificant. Both the percentage of one bedroom, PONEB, and the percentage of three or more bed rooms, PTHREEB, are positively and significantly related to maintenance costs. The input price variable, HOUSEWG, as in the laundry and linen equation is positively and significantly related to costs. ICF homes have significantly lower costs than either mixed or SNF facilities but differences between mixed and SNF facilities are not apparent. The payment mechanism variables are the only other significant regulatory variables in the equation. With regard to ownership, non-profit non-government homes appear slightly more expensive than for-profit and government homes.

Table T-4.C-4: DETERMINANTS OF PER DIEM LAUNDRY AND LINEN COSTS; NATIONAL NCHS DATA

Explanatory Variable	Estimated Coefficient	Standard Error
Output: Scale		
CONSTANT	.710	
AD1	-.048xxx	.008
AD2	.005x	.001
AD3	.0008x	.0004
AD4	.0006	.0006
AD5	.00007	.0007
AD6	-.00001	.0002
OC1	.000008	.001
OC2	.0619xx	.008
OC3	-.013	.009
OC4	-.006	.010
OC5	.008	.006
OC6	-.009	.007
Output: Patient Mix		
AGE	.002x	.0014
FEMALE	-.0008xx	.0004
Diagnosis		
ARTERY	.0005	.0003
OLD	.001xx	.0004
DIGEST	.002	.001
SENSE	-.008	.0007
Mental Status		
NERVOUS	-.0006xxx	.0002
SENILE	-.0003	.0003
Activities of Daily Living		
GROOMPD	.0009xxx	
GROOMD	.0008xxx	-.0003
DRESSD	.0007x	.0004
WALKPD	-.0008xx	.0004
WALKD	-.0006	.0005
BEDPD	.0006xx	.0003
BEDD	-.0008	.0008
Output: Case Flow		
ADMISPD	4.108xxx	1.511
Input Prices		
RETAILWG	.022	.015
HOUSEWG	.093xxx	.005
SMSA	.019	.014

Table T-4.C-4: DETERMINANTS OF PER DIEM LAUNDRY AND LINEN COSTS; (Cont'd.) NATIONAL NCHS DATA

Explanatory Variable	Estimated Coefficient	Standard Error
Regulatory Variables		
ICF	-.036xx	.017
ICF-SNF	.029x	.015
PM1	-.044xxx	.016
PM2	-.049xx	.016
CB	.021	.014
Facility Characteristics		
CHANGE	.061	.048
PCONTRACT	.003x	.001
R^2	.267	
Adjusted R^2	.243	
F-ratio	10.124	
N	1127	
Standard error	.206	

x significant at .10 level
xx significant at .05 level
xxx significant at .01 level

Table T-4.C-5: DETERMINANTS OF PER DIEM MAINTENANCE COSTS
NATIONAL NCHS DATA

Explanatory Variable	Estimated Coefficient	Standard Error
Output: Scale		
CONSTANT	1.025	
AD1	-.015	.011
AD2	.001	.012
AD3	.002xxx	.0005
AD4	.00005	.0007
AD5	-.0005	.0009
AD6	-.0005	.0003
OC1	.006xx	.001
OC2	.008	.010
OC3	.007	.011
OC4	-.030xxx	.013
OC5	.010	.007
OC6	-.0003	.008
Output: Patient Mix		
AGE	.0002	.001
Diagnosis		
ARTERY	.0004	.0004
STROKE	.002x	.0007
ACCIDENT	.001	.001
DIGEST	.005xx	.001
OTHARTERY	-.001	.001
SKELETAL	-.001	.0009
Mental Status		
SENILE	.0009xxx	.0003
SLEEP	.0006	.0004
Activities of Daily Living		
GROOMPD	.0008	.0005
DRESSPD	-.002xxx	.0006
DRESSD	.0008	.0005
WALKPD	.0006	.0004
BEDD	-.001	.001
Output: Services		
PONEB	.002x	.0005
PTHREEB	.002x	.0008
Output: Case Flow		
ADMISPD		
DEATH		
Input Prices		
RETAILWG	.030	.019
HOUSEWG	.092xxx	.007
SMSA	-.054x	.018

Table T-4.C-5: DETERMINANTS OF PER DIEM MAINTENANCE COSTS
(Cont'd.) NATIONAL NCHS DATA

Explanatory Variable	Estimated Coefficient	Standard Error
Regulatory Variables		
ICF	-.040x	.021
PM1	-.037x	.021
PM2	-.048x	.026
LIM	.041	.025
CB	.019	.019
OCCUPEN	.027	.022
NORTHEAST	.1216xxx	.028
CENTRAL	-.023	.028
SOUTH	.047	.029
Facility Characteristics		
NONPROFIT	.052*	.025
R^2	.263	
Adjusted R^2	.235	
F-ratio	9.453	
N	1127	
Standard error	.269	

x significant at .10 level
xx significant at .05 level
xxx significant at .01 level

T-4.C.3 Summary of Cost Center Analyses

To summarize the econometric results of this section, with regard to patient days, there is an approximate non-linear saucer-shaped curve for the payroll and nursing cost departments indicating the presence of slight economies of scale. In addition, it appears that there are slight economies of scale for food and laundry and linen as well.

No significant consistent relationship is evident between occupancy rates and components of costs, although isolated significant regression coefficients appear in the regresssion table.

The admission rate (i.e., the number of admissions per patient day) was allowed to enter the cost component functions of payroll expenses, nursing expenses, and laundry and linen. A strong statistical relationship exists between the admission rate and each of the three cost components. Examining the coefficients of this variable across the models reveals that the total cost of an admission (approximately $415) is composed primarily of payroll expenses (about $297) of which about 47 percent ($142) is nursing expenses. The laundry and linen department incurs an admissions cost of approximately $4.10.

It is not surprising that case mix affects the five cost components in different ways as suggested by the presence of different case-mix variables across the models. The average patient age appears as a significant explanatory variable in only the laundry and linear cost function. No consistent significant relationship was found between service intensity and the various cost components.

A strong positive statistical relationship exists between each cost category and those variables which proxy input prices. For each department, the corresponding wage rate (i.e., the average nurse's wage or cook's wage, the average dietitians or the maintenance personnel wage for the laundry and linen, and maintenance models) had the highest statistical significance.

Certification status has strong cost implications for the cost components of payroll expenses and nursing expenses; for the food, laundry, and maintenance categories, there are small but significant differences in costs on the basis of certification. For all categories, though, ICF status implies lower costs, all else constant.

With regard to regulation, a consistent finding across the cost components was that cost related reimbursement (either reasonable, actual, or retrospective) results in lower component costs per patient day than prospective reimbursement. By cost component category, the cost differential between the two types of systems ranged from a high of $1.38 for payroll to a low of $0.04 for maintenance. The cost differential is significant in all but the food and dietary category. In addition, flat rate systems are less costly then prospective systems for all of the cost component categories. Among these categories this difference is significant for all but the food and dietary category.

Regulations maintaining that a RN or LPN be present irrespective of the size of the home has a positive relationship to payroll costs and nursing costs. The enforcement of life safety codes appears to be significantly related to the food and dietary component. The requirement that a uniform chart of accounts be maintained and the time of patient assessment relative to the time of admission are cost determinants for the components of payroll and nursing; however, the relationships are significant for only the payroll category. The regional dummies are significant explanatory variables for payroll and nursing expenses, but there is no definite pattern to their significance for the remaining cost categories.

Lastly, for-profit homes are less costly than non-profit homes in all but the laundry and linen component. It is interesting to compare the magnitude of the cost differential across departments; the cost differential is highest on payroll expenses (about $1.48 per patient day) and lowest in the maintenance category (only $0.05 per patient day).

TECHNICAL PAPER #5

Technical Paper #5
NURSING HOME COSTS IN NEW YORK*

T-5.0 Introduction

This is the first of three technical papers reporting separate econometric studies of nursing home operating costs in each of three states--New York, Massachusetts and Indiana. This paper reports on the New York study; Technical Papers #6 and #7 report the Massachusetts and Indiana studies. These state-specific analyses were meant to supplement the analyses from nationwide data, reported in Technical Paper #4, and, in particular, to furnish "benchmarks" for assessing the extent to which the national results generalize to individual states.

In New York data came from the Medicaid cost reports submitted by long term care facilities in each of three consecutive years--1974, 1975 and 1976. New York was selected as one of the states to be analyzed primarily because the cost reports for New York state, especially 1975 and 1976, require greater reporting of services and patient characteristics than is available from the cost reports in most other states. That is, the New York reports were a "rich" source of relevant data. Furthermore, the New York cost reports were available in machine readable form and, most importantly, access to data was authorized by the New York State Department of Health. However, the data, and therefore our analyses from it may not generalize very well to the future for the following reason.

In early 1976 the so-called Moreland Commission--the New York State Moreland Commission on Nursing Homes and Residential Faciities--completed an intensive investigation of long-term care and long-term care regulation in New York. Their summary report (New York, 1975) concluded that "poor care" and "excessive profits" were serious problems in New York nursing homes and that regulatory requirements had been circumvented due to "bureaucratic ineptitude" and "tax enforcement." In consequence of the Moreland Commission report and other public revelations of scandal in the New York nursing home system, the New York regulatory authority has been reorganized and both regulation and enforcement are alleged to have been tightened. Thus, if the allegations of the Moreland Commission were substantially correct and these

---

* by A. James Lee and Howard Birnbaum

problems are now being corrected, our analysis of historical costs in New York may not generalize reliably to the future. Nevertheless, it is not altogether apparent that such problems are any more severe in New York than in other states.

Long-term care facilities in New York have been reimbursed on a prospective basis since 1971. Roughly speaking, the rate has been set in advance as the lower of (1) the actual inflation adjusted per diem cost in an earlier year and (2) a cost ceiling computed as the 60th percentile weighted average for the group. Prior to 1976 facilities were grouped by type (SNF vs. HRF), sponsor (non-profit vs. profit vs. public), bed size (0-49, 50-99, 100-199, 200-299, and 300+) and region (West, Rochester, Central, Northwest, North Metro, Long Island and New York City), except that groups with few cases were combined with others.[1] In 1976, sponsorship was eliminated as a grouping category. However, a new provision for waiver of the cost ceiling may well have blunted the impact of this change on the higher cost non-profit and public facilities. Mid-term rate adjustments were also eliminated in 1976. Because of such rate setting and other regulatory activity in the New York nursing home industry, including the non-statutory changes precipitated by the Moreland Commission, it will not be possible to reliably discriminate the effects of particular changes in the reimbursement mechanism using only three years of data. Although year-to-year differences in cost behavior can be distinguished, it is not possible to associate those differences with the many reimbursement/regulatory and possibly other relevant factors (e.g., the energy crisis) that also varied concurrently. The time series is simply too short and the regulatory situation in New York has been too much in flux. (A more detailed discussion of the New York prospective reimbursement system and its evolution can be found in Appendix T-5)

The remainder of this chapter is organized as follows. Section T-5.1 briefly summarizes the conceptual model developed more fully in Technical Paper #3; Section T-5.2 discusses the data and presents the operational variable specifications for analysis; Section T-5.3 presents the results from

---

[1] SNF is a Skilled Nursing Facility and an HRF is a Health Related Facility, essentially the same as an ICF or Intermediate Care Facility in other states.

a cross-sectional analysis for each of three years (1974, 1975 and 1976); Section T-5.4 presents results from certain first-difference analyses; and, finally, Section T-5.5 summarizes the results and explores the policy implications.

T-5.1 Review of the Conceptual Perspective

This section is meant only to review and, in some degree, recast the conceptual or theoretical model of nursing home cost behavior developed in Technical Paper #4. The reader should refer back to that paper if a fuller discussion is desired.

The primary interest in this paper, and in the two that follow, will be to specify and estimate cost functions for nursing homes within specific states. Whereas the conventional derivation of a cost function from economic theory indicates that cost is correctly a function of the output level or quantity (Q) and the relative prices of inputs (P_1, P_2, ..., P_n), as shown in T-5-1--the cost function is considerably more complicated in the case of the nursing home industry.

$$(T\text{-}5\text{-}1) \qquad C = f(Q; P_1, P_2, \ldots, P_n)$$

First, the "output" or "product" of nursing homes cannot be meaningfully measured as a single quantity. Whereas patient days of care is a useful indicator of the output level, it is only one of many dimensions describing relevant differences between the outputs of different facilities. In particular, the number of patient days per se reflects neither variation in the "quality" of care nor variation in the characteristics of patients receiving care--e.g., it is likely to be more expensive, other things equal, to care for a non-ambulatory patient than for an ambulatory patient. Thus, the level of output, Q, in the nursing home cost function is properly a vector of characteristics (Q_1, Q_2, ..., Q_m), including the number of patient days but also including patient, service, and quality characteristics.

Second, the nursing home industry is a regulated industry. If, for example, the nursing home industry had been unregulated there would be no reason to distinguish facility type (SNF vs. HRF) since the relevant differences would be captured by patient and service variables. However, the state and federal governments have imposed differential staffing and other requirements on different facility types and, to the extent that these regulations

do not fully reflect underlying resident requirements, the differential regulatory requirements may imply additional differences in cost by facility type.

In addition, the use of region, sponsorship, or bed-size categories in grouping facilities for rate setting means that historic differences in nursing home practice patterns (e.g., the "amenity" level) will be maintained as a consequence of regulation, independent of patient requirements, "quality" or regional variation in wage levels.[1] Finally, public financing of long-term care may impact costs to the extent that Medicaid and Medicare provide either more or less generous reimbursement for the same level and quality of care. Alternatively, private patients, on average, may "demand" a higher amenity level of care than can be provided to public patients within the ceiling limitations for reimbursement.

A third more general point is the following. One may reasonably choose to have the cost function reflect alternative techniques for producing the same output. With respect to nursing homes this can mean cost differences related to facility differences (e.g., age of facility) or other differences in operation of the facility (e.g., percent of services contracted).

In summary of the above discussion, we believe that the following is a more accurate conceptual representation of the nursing home cost function:

$$(T\text{-}5\text{-}2) \qquad C = f(Q_1, Q_2, \ldots, Q_m; P_1, P_2, \ldots, P_n; R_1, R_2, \ldots, R_x; F_1, F_2, \ldots, F_y)$$

where C is average per diem operating costs,

$Q_1, Q_2, \ldots, Q_n$ are quantity dimensions (e.g., patient or service characteristics),

$P_1, P_2, \ldots, P_n$ are the input prices (e.g., regional wage level),

$R_1, R_2, \ldots, R_x$ are variables reflecting the impact of regulation (e.g., grouping categories, or extent of governmental financing, and

---

[1] These variables might have once been used to proxy systematic patient and wage level differences. However, with the genesis of more detailed reporting requirements--such as in New York--it becomes unnecessary to rely upon such proxy measures for taking account of relevant differences; they can be measured directly and included explicitly in a reimbursement format.

F_1, F_2, ..., F_y are variables denoting differerences in technology or production processes (age of facility).

This expression gives the determinants of long-run average costs. However, in actual estimation the occupancy level, a measure of capacity utilization, should also be included to take account of short-run disequilibrium.

T-5.2 Empirical Specifications for Estimating Average Operating Cost Functions

This section indicates the operational specifications for estimating nursing home average operating cost functions, from information available on the New York Medicaid cost reports for 1974, 1975 and 1976. As will be seen, the information available is somewhat different in each of these years. In particular, only very limited data were available from the 1974 reports. Because of such problems we did not emphasize pooled cross-sectional/time-series analysis. The common variable list is too small. Instead we conducted separate cross-sectional analyses from the data available for each of the three time periods. However, in addition to estimating these cost functions, we have also conducted analyses of the year-to-year changes or first-differences in average operating cost.

The dependent variable for purposes of estimating the cost functions is the average or per diem operating expense, as it was in cost modeling from national data in Technical Paper #4. It is called AOC for Average Operating Cost.

We now discuss the independent variables that were available from one or more of the three years of data and comment briefly on the hypotheses associated with each. The reader will note that many of the variables and associated hypotheses are the same as those indicated in Technical Paper #4 in the analysis of the national data. Whereas our conceptual approach has suggested "novel" interpretations for several variables--e.g., that sponsorship is a regulatory-type variable in the state of New York--we group the variables here in a more conventional fashion and subsequently note their correspondence to the conceptual categories in equation (T-5-2). At any rate, we group the variables into the following major categories.

I. Facility Variables

- scale
- occupancy
- sponsorship (for-profit, non-profit or public)
- certification (SNF, HRF or SNF-HRF)
- percent of patients in one-, two- and three-bed rooms
- age of facility
- extent of contracting
- survey certification ratings
- wage level

II. Patient Variables

- age
- sex
- debility (e.g., percent non-ambulatory)
- payor source
- length of stay (or admission rate)
- percent deaths

III. Service

- services offered
- intensity of service use

IV. Regional Variables

- facility's location
- regional wage level
- population density

Facility Variables. Although average daily census might be a more meaningful indicator for the overall scale or level of output--as noted in Technical Paper #4--we nevertheless shall use the number of beds (BEDS) to proxy the scale of operations. The reason is as follows. In the state of New York five bed size groups--0-49 beds, 50-99 beds, 100-199 beds, 200-299 beds and 300 or more beds--were used as grouping categories for rate setting. We wish to incorporate these bed-size categories directly into the analysis since they are regulatory-type variables. In addition, we also wish to indicate non-categorical or continuous scale variables in order to test for economies or diseconomies of scale. The bed-size categorical variables may be picking up two effects--regulatory and scale. This will be explained at

length in our discussion of the results. In any event, we have chosen to use the number of beds instead of average daily census for the scale factor so that cost variation can be assessed in light of the New York grouping mechanism.[1]

Percentage occupancy (OCC) will be included as a measure of capacity utilization and, as such, will be thought to control for shorter-run or disequilibrium deviations from a longer-run cost curve. As is usual in economic analyses we hypothesize that as firms operate closer to capacity, production is more efficient and costs are lower.

Sponsorship--for-profit (PROFIT), non-profit (NONPROFIT) and public (the omitted category)--will be regarded as regulatory-type variables since facilities were grouped by sponsorship type for rate setting purposes until 1976. Facility certification--skilled nursing facility (SNF) or health-related facility (HRF)--was used for grouping in all years and thus serves a similar purpose. In addition, certification reflects the cost impact of those certification requirements imposed regardless of whether or not the required differences are needed for the actual resident population. For analytic purposes combined facilities, i.e., facilities having both SNF and HRF parts, were split into their SNF or HRF components.[2] Thus, we note two additional certification related differences, whether or not a SNF is part of a combined facility (SNF-HRF) and whether or not an HRF is part of a combined facility (SNF-HRF). However, the interpretation of these certification variables is considerably different from our interpretation of the SNF vs. HRF distinction. The SNF-HRF and HRF-SNF variables are thought to reflect economies (or diseconomies) due to "joint production" and to control for possible arbitrage in the accounting of costs between facilities--e.g., systematic allocation of SNF costs to the HRF part.

---

[1] Bed size and average daily census correlate so highly ($r^2 = .955$) that a methodological distinction between them is rarely meaningful, regression results using both are virtually equivalent.

[2] The data were reported in this fashion for 1974 and 1975 and could more easily be put in this form for 1976 than to aggregate components for 1974 and 1975.

The ratings from survey certification reflect the extent to which certification requirements have been met and thus regression coefficients on these variables indicate the extent to which costs associated with certification have been borne by the home. Of course, to the extent that certification assures the "quality" of care, these rating variables may also correspond to quality differences within the facility.[1] The ratings and their distribution were as follows:

UN	--	Unsatisfactory	(0.2%)
NI	--	Needs Improvement	(7.6%)
GF	--	Good, Federal	(29.1%)
GS	--	Good, State	(55.4%)
VG	--	Very Good	(0.2%)
DF	--	Deferred	(5.7%)
RNF	--	Deferred, New Facility	(0.2%)
RO	--	Deferred, New Owner	(0.4%)

The figures in parentheses give the relative frequency of each designation among observations with non-missing data--84.8 percent of the total sample size. Except for the deferred categories (DF, RNF, and RO) the designations are displayed in ascending order of "satisfactoriness." Unlike most states, the state certification requirements in New York are more stringent than the federal requirements. The circumstances under which the DF designation was given are not known. Furthermore, the less than satisfactory NI and UN ratings may be associated with altogether different deficiencies in different

[1]Although the final ratings were available for 1976 we chose not to use them, both for theoretical and empirical reasons. First, the final ratings, as distinguished from the initial ratings, reflect considerable negotiation, including negotiation of improvements or changes to be implemented in the future. Thus, the initial ratings should reflect more fairly the situation that prevailed during the course of the year. This was borne out by preliminary analyses that explored both alternatives. Only 0.8 percent of facilities were designated as needing improvement in the final overall rating. In the initial rating 7.6 percent of facilities received this designation and 0.2 percent were judged to be altogether unsatisfactory. Thus, we have used the initial ratings in the results reported here. These data were provided in "hard copy" by the New York State Department of Health.

facilities--e.g., cleanliness in one, nursing in another and food in a third. After exploring several alternatives and taking account of designations with only a few cases, the rating designations were aggregated into the following categories: "good" (GS, GF or VG), "needs improvement" (NI or UN), and "deferred" (DF, RNF or RO). Preliminary analyses found no meaningful difference in costs between facilities receiving either the GS or GF designation. Since these ratings--good (the omitted category), needs improvement (NI) and deferred (DEFERRED)--were the only explicit measures of quality available. It is especially important to include them and also to use the 1976 ratings in the 1974 and 1975 analyses. This assumes that deficiences noted in 1976 were not of recnt origin but had persisted over a period of time. The consistency of the findings with respect to these variables in all three years lends support to this assumption.

The percentages of patients in one-(PCT1B)-two (the omitted category), and three or more (PCT3B) bed rooms measure differences along a dimension of "quality", namely, the degree of privacy among patients. These variables may also reflect ddifferences in the technology or process of producing the same outputs. The percentage of operating expense contracted (%CONTRACT) is included as a control for a potentially different production process. In theory, one would not contract for services unless they cost less than the alternative of having the facility produce them itself. If so, facilities with a larger percentage of contracted expense should have lower operating costs. Pragmatically, the situation may be very different, however, inasmuch as contracting is alleged to be a common mechanism for the fraudulent overcharging of publicly-reimbursed patients--e.g., via "kick-backs" from the ccontractors. The age of the facility (FACILITY AGE) is included to control for the cost implications of potential "vintage" effects in the substitution of capital and labor which result in exogenous differences in the production process available to operators. Finally, there should be for each facility a vector of relative prices for the various inputs (e.g., nurses, food and equipment) used in production of long-term care. Whereas some regional proxy variables are available--as will be seen--there is little explicit information on the relevant price differences between facilities. For 1975 we calculated the average LPN wage (LPN WAGE) for each facility. However, the requisite data were not available for 1974 and 1976.

Patient Variables. The age distribution of residents or patients is typically included in nursing home studies to proxy age-related differences in infirmity and the consequent "need" for differential treatment and resources. The usual hypothesis is that the cost of caring for an older patient is higher, on the average, because older persons tend to be more disabled and require more personal attention. Of course, to the extent that relevant differences in debility are measured more explicity (e.g., percent non-ambulatory) the age proxies may no longer be meaningful. In any event, we included the percentage distribution of patients aged 64 and under (%UNDER65), aged 65 to 74 (the omitted category), age 75 to 84 (%AGE75TO84) and over 84 (%OVER84) to test for any uncontrolled effects of patient age.

AMS (1976) and other studies have indicated that female patients are more expensive. We explored this prospect by including percent female (%FEMALE) as a patient variable.

The New York cost reports for 1975 and especially 1976 have fairly extensive information on patient debility levels. For both 1975 and 1976 the reports have a broad range of information on patient functional status--the number of patients requiring no assistance, the number requiring partial assistance, and the number requiring full assistance with each of the following: (1) walking, (2) eating, (3) bathing, (4) breathing, (5) dressing, (6) correspondence, (7) medication and (8) massage.[1] (This information is available only for walking in 1974.) For analytic purposes we have computed the percentages of patients requiring either some or full assistance with each of these (%WALK, %EAT, %BATHE, %BREATHE, %DRESS, %CORRESPOND, %MEDICATE and %MASSAGE). In addition, for 1976 we have information on patient impairment--the number of patients having no impairment, the number having partial impairment and the number having total impairment--along each of the following dimensions: (1) alertness, (2) sight, (3) hearing, (4) speech, and (5) communications. We have analogously computed for analytic purposes the

[1]Unfortunately, the data on medication was not usable in 1975.

percentage of patients having either partial or total impairment (%ALERT, %SEE, %HEAR, %SPEAK, AND %COMMUNICATE).[1], [2]

The percentage of patients dying each year (%DEATHS) calculated as the number of deaths divided by the average daily census, might be a further indicator of patient debility or health status. On the other hand, it may also reflect, at least in some degree, the unfortunate consequences of lower quality care. Although this variable is included, its interpretation remains ambiguous pending further study.

As seen in Technical Paper #4, the average length-of-stay and the admission rate are very nearly non-linear transformations of one another, given other variables already included in the model. We use an admission rate variable (ADMRATE)[3] computed as the total number of admissions divided by total patient days. This definition yields a coefficient that is a straightforward estimate of the (average) cost per admission.

---

[1]The absence of data on mental health status is conspicuous.

[2]Although we might have indicated a separate variable for partial and total impairment and for the requirement of partial vs. full assistance, an already serious multi-collinearity problem would have been much aggravated. Of course, one alternative would have been to indicate the percentage of patients requiring either full assistance only or having only total impairment. However, preliminary analyses revealed that the more inclusive alternative which was chosen gave superior statistical results. Still another alternative might have been to use all the data in constructing an index of patient "costliness." The usual indices of debility may not be appropriate for cost modeling purposes. Certain dimensions of debility may be more or less costly than others, or even conceivably cost less than for otherwise similar persons without the impairment (e.g., a patient who is not alert might require less recreational therapy). The construction of a more appropriate index is certainly worthwhile, and might well make the resultant reimbursement formula more meaningful to some, it was not thought to be within the scope of this project. If it is to be done well, as it must be if it were to be used in a reimbursement formula, it would be a major task.

[3]There is reasonably good information on the distribution of stays from the New York cost reports and a computed average length-of-stay variable was indicated in preliminary analyses. However, the admission rate variable (ADMRATE) gave vastly superior statistical performance.

The last patient variables available from the cost reports are the payor characteristics. We have calculated the percentage distribution of private (%PRIVATE), Medicaid (the omitted category) and Medicare (%MEDICARE) for each facility. As discussed in Section T-5.1, we regard these as regulatory-type variables which indicate the extent to which Medicare and Medicaid are systematically purchasing more or less expensive care than private patients. (The difference in costliness is measured after controlling for differences in certification, "quality," and patient characteristics.) Of course, any findings cannot indicate whether or not the cost differences correspond to "real" differences in the care received by patients or whether they simply reflect inefficient production or excessive reimbursement.

Service Variables. Service variables are basically regarded as proxying "quality" differences. We have information from 1975 and 1976 on service intensity and have computed the following variables for analysis: annual number of pharmacy prescriptions per patient (#PRESCRIPTIONS); annual number of x-rays per patient (#X-RAYS); annual number of physical therapy sessions per patient (#PT); annual number of speech therapy sessions per patient (#SPEECH THERAPY); annual number of occupational therapy sessions per patient (#OCC THERAPY); annual number of inhalation therapy sessions per patient (#INHAL THERAPY); annual number of radiation therapy sessions per patient (#RADIATION THEPAY); and annual number of social service sessions per patient (#SOC SERVICE).[1] Facilities providing greater service intensity are presumed to be using more resources and thus are hypothesized to be more expensive. However, the relationship between service intensity and "quality" is a tenuous one at best. Our approach cannot distinguish whether or not the services provided are appropriate, i.e., too few or too many given the patient mix of the home. The usual presumption has been that the more services, the better. Whereas this may be a reasonable assumption for the overall average across facilities, it is probably not true with respect to selected facilities. Unfortunately, more meaningful information on "quality" differences was not available.

---

[1] A further service intensity variable, the annual number of activity therapy sessions per patient, was omitted because of its collinearity (r^2=.88) with #SOC SERVICE. The annual number of tests per patient was also omitted for analogous reasons.

A further variable available in the same series was the annual number of utilization reviews per patient (#UR). Although it may be considered a service to patients by insuring that an appropriate mix of services are provided, it is also a regulatory-type variable. It was available for 1976 only.

For 1975 only, the cost reports identified those services provided to Medicaid patients that were included in determining the Medicaid rate. Although these variables overlap the service intensity variables to some degree, they serve a somewhat different function. They control for the accounting of service expenses indicating whether or not the resource cost is reflected in the rate. Categorical (0-1) variables indicating whether or not the service was provided to Medicaid patients and its cost included in the rate were defined for the following: drugs (DRUG); physical therapy (PHYS THERAPY); occupational therapy (OCC THERAPY); speech pathology (SPEECH PATH); special duty nurses (SPECIAL DUTY); and physical services (PHYS SERV). Although similar information was available for (i) dental services, (ii) oxygen, (iii) audiology and (iv) podiatry, we judged a priori that these service components were unlikely to be costly enough to warrant indicating them as separate variables for analysis. Moreover, very few homes provide these services. We have aggregated these services to form an index (OTHER SERVICES), ranging between zero and four, indicating the number of such services provided to Medicaid patients and included in the rate.

Regional Variables. Among the regional variables are categorical, zero-one variables (REGION1 to REGION7) indicating the seven regions--West, Rochester, Central, Northwest, North Metro, Long Island and New York City--used in grouping facilities for reimbursement. (All seven regions were used in 1974 and 1975; six regions were used in 1976. The North Metro and Long Island regions were combined in that year.) Since explicit measures of regional wage and price differences are included in the cost models, these location dummies viewed as reflecting regional practice differences that are possibly maintained by operators due to having facilities grouped by region for Medicaid rate setting purposes. That is, they are essentially regulatory-type variables.

In addition to the facility-specific LPN wage in 1975, we also include the average wage (RETAIL WAGE) for retail workers in the county of facility location as a proxy for regional variation in wage and other price levels. The county-wide population density (POPDEN) is also included for this purpose, under the assumption that prices tend to be higher in more populated areas. This variable may, however, also reflect agglomeration economies associated with urban residence--e.g., access to specialized services such as contract laundries. These data come from the HRA Area Resource File (1977 version).

Parameter Specification

Considered here is the parametric form for estimation of the cost models. There is virtually an unlimited number of ways in which the data might be analyzed, both with respect to the parametric forms to be estimated and the variables to be included. Because of the potential complexity, expense and intractability of reporting a broad range of alternative parametric and variable specifications, we have pursued the following reasonably general empirical strategy.

First, the parametric form estimated should be amenable to manipulation for formula reimbursement. Whereas more sophisticated users would have no difficulty with non-linear alternatives, a formula developed from a non-linear relationship might be considerably less intuitive for state rate setters and nursing home operators. Furthermore, in the absence of prior information on the functional form it has now become standard to estimate a linear approximation to the "true," but unknown, functional form. (The reader also should recall that exploratory investigations of log-linear alternatives in the national analysis (see Technical Paper #4) did not yield a superior statistical model--i.e., the linear model gave a better fit.)

However, the model estimated here is not strictly linear with respect to a few "key" variables. Much previous cost modeling, both with respect for nursing homes and for hospitals, has been concerned with economies and diseconomies of scale and with potential non-linearities with respect to occupancy. We did not wish to ignore these interests and yet we did not wish, as others have done, to impose an arbitrary quadratic form on the relationship of these variables. We chose instead to use a piece-wise linear estimation technique, also called "grafting." This procedure, which has been

largely ignored in the health care literature, is general in the sense that it is empirically-based and permits the data to indicate a piece-wise linear approximation to an arbitrary but unknown functional form.[1] This is accomplished by estimating separate linear segments over different ranges of the independent variable with the constraint that the relationship is continuous. See Figure T-5.1(a). For example, if the independent variable in that figure were beds and the dependent variable average operating cost, the results would indicate economies of scale between 0 and X_1, no economies between X_1 and X_2, some diseconomies between X_2 and X_3, and substantial diseconomies above X_3. It is also possible to allow for discontinuities in the relationship as shown in Figure T-5.1(b).

For occupancy we estimated a piece-wise relationship of the kind shown in (a). In particular, we obtain separate slope coefficents for each of the following percentage occupancy ranges: (i) 0-80 percent, (ii) 80-90 percent, (iii) 90-95 percent and (iv) 95 percent and higher. The estimation involved the specification of four separate occupancy variables (OCC1-OCC4) with each of them defined as shown in Table T-5.1.[2]

For our scale factor--the number of beds--we estimate a relationship of the type shown in Figure T-5.1(b). Our reasoning is as follows. Recall that five bed-size categories (0-49, 50-99, 100-199, 200-299 and 300+) are used as grouping criteria for rate-setting purposes. We obtain a type-(b) function if piece-wise (continuous) variables are defined over the same grouping bed-size ranges(BEDS1-BEDS5). As already noted, scale and regulatory effects are intertwined in the resultant relationship. We shall return to this issue in discussion of the results.

Although the independent variables were developed above in their more conventional groupings--(1) facility variables, (2) patient variables,

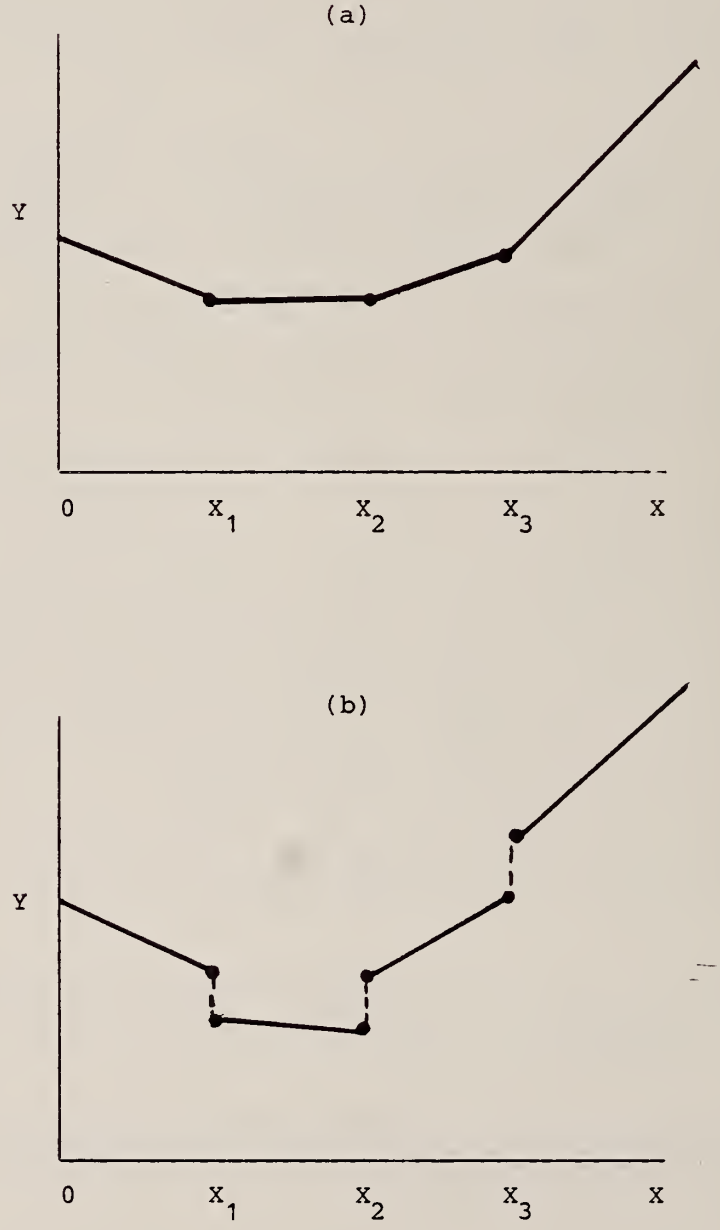
---

[1] Multi-collinearity is also a less serious concern with respect to this alternative. In particular, bed-size and bed-size squared--the usual specifications--correlate very highly ($r^2 > .9$).

[2] To take an example, if actual occupancy were 93 percent, then from the definitions in Table T-5.2 we would have that OCC1 equals 80, OCC2 equals 10, OCC3 equals 3, and OCC4 equals 0.

Figure T-5.1

ILLUSTRATION OF PIECE-WISE ESTIMATION



(3) service variables, and (4) regional variables--we have in Table T-5.2 reorganized them into their more conceptually-appropriate categories--output characteristics (Q), input price variables (P), regulatory-type variables (R), technology or production process variables (F) and indicators of short-run disequilibrium (E). Table T-5.2 also defines the variables. However, the reader may wish to return to the above discussion for additional clarification of variables and conceptual motivation of the variable organization.

As already noted, we are estimating linear relationships between average operating cost (AOC) and the Q- , P- , T- and E- vectors. That is, we are estimating relationships of the following form;

$$(T\text{-}5.3) \quad AOC = b_0 + B_Q \cdot Q + B_P \cdot P + B_R \cdot R + B_T \cdot T + B_E \cdot E + u,$$

where the B's are vectors of coefficients and u is the disturbance term.

The independent variable list is lengthy and it should be apparent that the potential for multi-collinearity is great. However, the reader is reminded that multi-collinearity merely leads to inefficiency of parameter estimates and does not affect the "goodness-of-fit." If the purpose were only to explain cost variation, using the available data, even severe multi-collinearity would not be a problem. However, instability of coefficient parameter estimates may be troublesome to users of any reimbursement formula developed from the regression results. Although we have been sensitive to the potential for multi-collinearity in specifying the model--especially with respect to the specifications for bed-size and occupancy--we see no fully satisfactory alternatives for resolving the difficulty. Sound conceptual rationale has been given for including each of the variables included in the model. In practice, the multi-collinearity problem is somewhat mitigated because all variables are not available in any of the three years for which we have data. However, the numbers of variables for 1976 and 1975 are quite large.

We have sought to further address the multi-collinearity concern by estimating not only complete models but also models using a stepwise selection of those variables having a t-statistic of one or greater--i.e., having a coefficient estimate larger than its standard error. (This corresponds to approximately a .30 confidence level.) Although stepwise regression procedures have not been popular of late, we invite the reader to examine the

Table T-5.2

SUMMARY OF VARIABLE DEFINITIONS FOR NEW YORK ANALYSES

OUTPUT CHARACTERISTICS

Scale (bed-size variables for piece-wise estimation)

BEDS1---------Bed-size segment defined over 0-49 bed interval, equal to the actual number of beds if 49 or fewer and (ii) equal to 49 if more than 49

BEDS2---------Bed-size segment defined over the 50-99 bed interval, equal to zero if actual number of beds is fewer than 50 equal to actual number of beds minus 49 if bed size ranges between 50 and 99, and equal to 49 if number of beds is more than 99

BEDS3---------Bed-size segment defined over the 100-199 bed interval, equal to zero if number of beds is fewer than 100, equal to number of beds minus 99 if bed size ranges between 100 and 199, and equal to 100 if number of beds is more than 199

BEDS4---------Bed-size segment defined over the 200-299 bed interval, equal to zero if number of beds is fewer than 200, equal to number of beds minus 199 if bed size ranges between 200 and 299 and equal to 100 if number of beds is more than 299

BEDS5---------Bed-size segment defined over the 300+ bed interval, equal to zero if number of beds is fewer than 300 and equal to number of beds minus 299 if bed size is more 300 or more

Patient Age Distribution
(percentage aged 65 to 74 omitted)

%UNDER65------Percent of patients under the age of 65

%AGE75TO84----Percent of patients aged 75 to 84

%OVER84-------Percent of patients over the age of 84

%FEMALE-------Percent female patients

Patient Functional Status Variables

%WALK---------Percent of patients requiring either partial or total assistance with walking

%EAT----------Percent of patients requiring either partial or total assistance with eating

%BATHE--------Percent of patients requiring either partial or total assistance with bathing

%BREATHE------Percent of patients requiring either partial or total assistance with breathing
%DRESS--------Percent of patients requiring either partial or total assistance with dressing
%CORRESPOND---Percent of patients requiring either partial or total assistance with correspondence
%MEDICATE-----Percent of patients requiring either partial or total assistance with medication
%MASSAGE------Percent of patients requiring either partial or total assistance with massage

Patient Impairment Variables

%ALERT--------Percent of patients either partially or totally non-alert
%SEE----------Percent of patients having either partial or total impairment of sight
%HEAR---------Percent of patients having either partial or total impairment of hearing
%SPEAK--------Percent of patients having either partial or total impairment of speech
%COMMUNICATE--Percent of patients having either partial or total impairment of communication

%DEATH----------Annual death rate as a percentage of average daily census

ADMRATE---------Ratio of total annual admissions to total annual patient days

Service Offerings--Dummy variables indicating:

DRUG----------Cost of drugs is included in the rate

PHYS THERAPY--Physical therapy is provided to Medicaid patients and its cost is included in the rate

OCC THERAPY---Occupational therapy is provided to Medicaid patients and its cost is included in the rate

SPEECH
THERAPY-------Speech therapy is provided to Medicaid patients and its cost is included in the rate

SPECIAL DUTY--Special duty nursing is provided to Medicaid patients and its cost is included in the rate

PHYS SERVICES-Physical services are provided to Medicaid patients and its cost is included in the rate

Table T-5.2 (Continued)

OTHER
SERVICES------An index measure of other service offerings equal to the number of the following services, provided to Medicaid patients and included in the rate: (i) dental services, (ii) oxygen, (iii) audiology and, (iv) podiatry

Service Intensity

#PRESCRIP-
TIONS--------Average annual number of pharmacy prescriptions per patient
#X-RAYS-------Average annual number of x-rays per patient

#PHYSICAL
THERAPY------Number of patients receiving physical therapy

#SPEECH
THERAPY------Number of patients receiving speech therapy

#OCC THERAPY--Number of patients receiving occupational therapy

#INHAL
THERAPY------Number of patients receiving inhalation therapy

#RADIATION
THERAPY------Number of patients receiving radiation therapy

#SOC
SERVICES-----Number of patients receiving social services

Bedroom Distribution (percentage of patients in 2-bed rooms omitted)

%1BED---------Percent of patients in 1-bed rooms

%3BED --------Percent of patients in 3-or more bed rooms

INPUT PRICE VARIABLES

LPN WAGE------Average LPN wage within facility

RETAIL WAGE---Average county-wide retail wages

POPDEN--------County-wide population density

REGULATORY-TYPE VARIABLES

HRF----------Dummy variable identifying Health Related Facilities (SNF omitted)

Table T-5.2 (Continued)

Sponsorship (public sponsorship omitted)

PROFIT--------Dummy variable identifying for-profit facilities

NONPROFIT-----Dummy variable identifying nonprofit facilities

Payor Distribution (percentage Medicaid omitted)

%PRIVATE------Percent private pay patients

%MEDICARE-----Percent Medicare patients

Survey Certification Ratings ("good" rating omitted)

NI------------Dummy variable indicating "needs improvement" rating

DEFERRED------Dummy variable indicating rating deferred

#UR----------------Number of utilization review cases divided by average daily census

Bed-Size Categories--Dummy variables indicating:

SIZE1
(omitted)-----0-49 beds

SIZE2---------50 to 99 beds

SIZE3---------100 to 199 beds

SIZE4---------200 to 299 beds

SIZE5---------300 or more beds

Regions--Dummy variables indicating:

REGION1
(omitted)-----West

REGION2-------Rochester

REGION3-------Central

REGION4-------Northeast

REGION5-------North Metropolitan

REGION6-------Long Island

REGION7-------New York City

Table T-5.2 (Continued)

TECHNOLOGY OR PRODUCTION PROCESS VARIABLES

SNF-HRF-------Dummy variables indicating SNF part of a combined facility

HRF-SNF-------Dummy variable indicating HRF part of a combined facility

CLINIC--------Dummy variable identifying facilities with a clinic

%CONTRACT-----Extent of contracting for services as a percentage of operating expense

FACILITY AGE--Age of facility, in years

INDICATORS OF SHORT-RUN DISEQUILIBRIUM

Occupancy Segments (used for piece-wise estimation)

OCC1----------Occupancy segment defined over the 0-80 percent occupancy interval, equal to actual occupancy if less than or equal to 80, and equal to 80 if occupancy is higher than 80 percent

OCC2----------Occupancy segment defined over 80-90 percent occupancy intervals equal to zero if occupancy is less than 80, equal to actual occupancy minus 80 if occupancy ranges between 80 and 90 percent, and equal to 10 if occupancy is equal to or higher than 90 percent

OCC3----------Occupancy segment defined over the 90-95 percent occupancy interval, equal to zero if occupancy is less than 90, equal to actual occupancy minus 90 if occupancy ranges, between 90 and 95 percent, and equal to 5 if occupancy exceeds 95 percent

OCC4----------Occupancy segment defined over the 95-100 percent occupancy interval, equal to zero if occupancy is less than 95 percent, equal to actual occupancy minus 95 if occupancy is greater than 95 percent.

results on their merit. As will be shown, what one learns from either model (full or limited) is basically the same. In both models, the same variables retain statistical significance and maintain their relative economic magnitude. Thus, the prospect of improper inference from the limited variable models is not serious.

T-5.3 The Cost-Function Results--Cross-Sectional Estimates for 1974, 1975 and 1976

This section reports and discusses results obtained from estimating the cost function models. Recall the allegations of the Moreland Commission--namely, that "poor care" and "excess profits" have been a problem in New York nursing homes--and their implications for interpretation of the results. If poor care is pervasive (i.e., not merely isolated cases) and this has not been captured by the proxies for "quality," it is possible that the cost-function models reported here under-estimate the cost of providing a minimally-acceptable quality of care. On the other hand, if excess profits are pervasive, the problem is just the opposite with the model overpredicting cost requirements. Unfortunately, this study cannot begin to assess the prevalence or quantitative importance of such prospects. Indeed such an assessment cannot be accomplished apart from having more explicit, facility-specific information on the incidence of such problems. In any event, it is not apparent a priori that such problems are any more or less severe in New York than elsewhere in the nation.

The data for each of the three years included all skilled nursing facilities (SNF's) and health related facilities (HRF's) submitting Medicaid Cost Reports in New York--637 facilities in 1974, 671 facilities in 1975 and 714 facilities in 1976. Recall that data for combined SNF-HRF facilities was partitioned and such facilities are represented twice, both as an SNF part and as an HRF part. All hospital-based SNF's and HRF's were omitted because of a concern that such facilities might be specializing in shorter-term care and thus not be comparable to free-standing facilities. The dependent variable--average operating cost (AOC)--was formed as the sum of all payroll,

dietary, laundry, housekeeping and maintenance expenses divided by total patient days.[1] Observations lacking any of this information were omitted. Several facilities with average operating costs less than $15 per day were also omitted. (Our analysis, in Section T-5.4, of the year-to-year cost changes indicated that such low cost levels were due to reporting error).[2] Furthermore, facilities not open for the full year were screened in 1975 and 1976, the only years for which the number of days open was available; such newly-opened facilities were much less likely to be operating on their long-run cost curves. For analogous reasons we sought to eliminate facilities changing bed-size over the year. Unfortunately, explicit information on bed-size changes was available for 1975 only. Thus, as an ad hoc alternative for 1974 and 1976, we also screened out facilities with average occupancy rates less than 60 percent.[3] The numbers of usable observations--i.e., observations that remained after screening--were 479 in 1974, 504 in 1975 and 497 in 1976. Missing data were not in general a serious problem.[4]

---

[1] The only facility costs excluded from operating expense were capital costs (depreciation, interest and rent), taxes and profit. The "true" capital costs can not be reliably measured from available accounting data; taxes depend crucially upon property value assessments that also tend to be unreliable; and profit is determined as a residual category and bears a very imperfect relation to actual profitability.

[2] Invariably, facilities reporting low cost levels in a given year were found to have extraordinary, and altogether implausible year-to-year cost changes.

[3] A few facilities reporting the patient's average age as being less than 40 were also omitted. It was not known whether such facilities actually serve a younger clientele or the patient age distribution was simply misreported.

[4] However, we must extend the following general caveat with respect to missing data problems. We can not be certain that all instances of missing data have been recognized as such. Missing value codes were not used in keypunching the cost reports and our software could not distinguish between a blank field and a "true" zero. Since consistency checks were not always possible, we may have unknowingly treated some missing values as if they were "true" zeros (e.g., %CONTRACT).

Mean values were substituted for missing values in some instances (NI, DF in all three years, %CONTRACT and LPN WAGE in 1975 and #UR in 1976). In addition, a few variables (%DEATHS in 1975, and %MEDICARE, %CONTRACT and FACILITY AGE in 1976) were omitted because the data were not useable, either due to errors in reporting or errors in data processing.

On balance, the 1975 model is the most fully specified of the three. Although the 1976 model also included many of the same variables, data problems were more severe in that year and we remain less confident about the quality of the 1976 data than for 1975. On the other hand, the 1976 results are more likely to reflect reforms engendered by the Moreland Commission. Some judgment must be exercised in weighing these competing factors. However, analyses of the cost data that become available in subsequent years will be needed to dependably resolve many of the specification issues, particularly if nursing home cost behavior in New York is changed due to the "reforms" now taking place.

As already noted, both "complete" and "limited" ($t>1$) cost models were estimated from each of the three years of data. The so-called complete or "fully specified" models are shown in Table T-5.3; the so-called limited or "partial" models are shown in Table T-5.4. We first discuss the complete model results.

T-5.3.1 The Complete or Fully Specified Model

These models include as explanatory variables all conceptually-appropriate variables that were available and useable in each of the three years. The results for different years may be different in part because the specifications are different. Since only a limited number of variables were available for 1974, we shall emphasize the more nearly comparable and comprehensive analyses from 1975 and 1976.

Beginning with the summary statistics, the coefficient of determination (R^2) rises from 64.7 percent in 1974 to 76.8 percent and 76.4 percent, respectively, in 1975 and 1976. The corrected coefficients of determination are, respectively, 62.2 percent, 73.6 percent, and 73.4 percent. The F-ratios are highly significant in each case. In 1974, 12 of 32 variables (or 38 percent) are significant at the .10 level or better. In 1975, 22 of 61 variables (or 36 percent) are likewise significant. And in 1976, 16 of 55 variables (or 29 percent) are significant at the .10 level or better. The

Table T-5.3

DETERMINANTS OF AVERAGE OPERATING COSTS:
COMPLETE OR FULLY-SPECIFIED COST MODELS, NEW YORK
(standard errors in parentheses)

Independent Variable	OLS Regression Coefficients 1974	1975	1976	Beta Coefficients 1975
		OUTPUT CHARACTERISTICS		
Scale				
BEDS1	0.0101 (.0465)	-0.00183 (.0598)	-0.102* (.0614)	-0.00104
BEDS2	0.00388 (.0195)	-0.0238 (.0269)	-0.0324 (.0215)	-0.0465
BEDS3	0.00271 (.0104)	0.00910 (.0101)	0.0118 (.0136)	0.0362
BEDS4	-0.0123 (.0216)	0.0146 (.0148)	0.00649 (.0145)	0.0369
BEDS5	0.00207 (.00645)	0.00364 (.00640)	-0.000977 (.00658)	0.0164
Patient Age Distribution				
%UNDER65	0.0456 (.0737)	-0.0101 (.0807)	0.00293 (.0813)	-0.00641
%AGE75TO84	0.0294 (.0418)	0.0114 (.0529)	0.0242 (.0499)	0.00986
%OVER84	0.0309 (.0375)	-0.0127 (.0448)	0.0292 (.0447)	-0.0164
%FEMALE	0.0412 (.0254)	0.0290 (.0296)	0.0451 (.0321)	0.0327
Patient Functional Status Variables (percentages of patients requiring assistance)				
%Walk	0.0334* (.0164)	-0.0119 (.0166)	0.0272 (.0168)	-0.0301
%Eat		0.0453***(.0134)	-0.00881 (.0134)	0.129
%Bathe		0.0319* (.0193)	0.0207 (.0179)	0.0587
%Breathe		-0.00634 (.00658)	-0.00144 (.00635)	-0.0258
%Dress		0.00550 (.0179)	0.0153 (.0164)	-0.0159
%Correspond		0.0229* (.0127)	0.00194 (.0114)	0.0720
%Medicate		-0.00258 (.0175)		-0.00404
%Massage		-0.00200 (.0102)	0.00788 (.0103)	-0.00636

*Significant at .10 level
**Significant at .05 level
***Significant at .01 level

Table T-5.3 (cont'd.)

Independent Variable	Regression Coefficients 1974	1975	1976	Beta Coefficients 1975
Patient Impairment Variables (percentages of patients with impairment)				
%ALERT			-.0159 (.0112)	
%SEE			.0204* (.0118)	
%HEAR			-.0100 (.0130)	
%SPEAK			-.00304 (.0153)	
%COMMUNICATE			-.00136 (.0142)	
%DEATHS	-.00625 (.0205)		-.0124 (.0254)	
ADMRATE	342* (191)	1050*** (148)	276 (273)	.332
Service Offerings (0-1)				
DRUGS		-.337 (.599)		-.0156
PHYS THERAPY		.211 (.852)		.00705
OCC THERAPY		.770 (.587)		.0364
SPEECH THERAPY		-.324 (.623)		-.0157
SPECIAL DUTY		-.667 (1.50)		-.0118
PHYS SERVICES		.935 (.654)		.0458
OTHER SERVICES		.520* (.300)		.0650
Service Intensity				
#PRESCRIPTIONS		.00132 (.00284)	.0121* (.00723)	.0114
#X-RAYS		.341*** (.112)	.613***(.101)	.0846
#PHYSICAL THERAPY		3.14*** (.976)	1.25* (.564)	.108
#SPEECH THERAPY		.343 (6.14)	13.1* (7.5)	.00163
#OCC THERAPY		.844 (1.47)	.151 (.336)	.0157
#INHAL THERAPY		14.6** (7.02)	-6.18 (12.6)	.0535
#RADIATION THERAPY		-.933 (3.83)	33.0 (30.0)	-.00582
#SOC SERVICE		1.02** (.429)	.437 (.407)	.104

*Significant at .10 level
**Significant at .05 level
***Significant at .01 level

Table T-5.3 (cont'd.)

Independent Variable	Regression Coefficients 1974	1975	1976	Beta Coefficients 1975	
Bedroom Distribution					
%1BED		.0316*** (.0158)	-.0131 (.0365)	.0715	
%3BED		-.0165 (.0118)	-.0312 (.0229)	-.0428	
		INPUT PRICE VARIABLES			
LPN WAGE		.701***(.187)		.106	.285
RETAIL WAGE	6.51* (2.12)	6.29***(2.31)	6.83* (2.30)	.179	
POPDEN	-.0000506 (.00003)	.00000865 (.00004)	.0000123 (.0004)	.0139	
Bed-Size Categories					
SIZE2					
SIZE3		1.12 (.782)		.0512	
SIZE4			-2.23* (1.10)		
SIZE5	4.89* (2.32)				
Regions					
REGION2	-.784 (1.19)	-1.36 (1.27)	-.759 (1.26)	-.0389	
REGION3	-.402 (1.02)	-.841 (1.11)	-.993 (1.20)	-.0281	
REGION4	-.629 (1.14)	-.696 (1.26)	-.888 (1.14)	-.0221	
REGION5	4.14* (1.50)	2.88* (1.58)	4.21*[a] (1.21)	.0965	
REGION6	4.92* (1.55)	5.07***(1.73)		.151	
REGION7	6.42* (1.53)	4.02**(1.76)	4.68* (1.58	.182	

*Significant at .10 level
**Significant at .05 level
***Significant at .01 level
[a]Regions were combined for grouping purposes in 1976.

Table T-5.3 (cont'd.)

Independent Variable	Regression Coefficients 1974	1975	1976	Beta Coefficients 1975
TECHNOLOGY OR PRODUCTION PROCESS VARIABLES				
SNF-HRF		-1.79**(.706)	-.973 (.722)	-.0745
HRF-SNF		-1.72 (1.14)	-.619 (1.14)	-.0691
CLINIC		-2.79 (1.97)		-.0342
%CONTRACT		-.0651 (.0514)		.0625
FACILITY AGE		-.0182 (.0130)		
INDICATORS OF SHORT-RUN DISEQUILIBRIUM				
Occupancy Segments				
OCC1	.119 (.140)	.0442 (.0428	-.0566 (.269)	.0484
OCC2	-.204**(.183)	-.223 (.179)	-.273 (.249)	-.0745
OCC3	.0845 (.288)	.267 (.300)	-.147 (.324)	.0507
OCC4	.184 (.185)	-.00893 (.102)	-.216 (.144)	.00265
CONSTANT	17.0	4.23	25.0	
Summary Statistics				
Coefficient of Determination (R^2)	.647	.768	.764	
Corrected Coefficient of Determination	.622	.736	.743	
F-Ratio	25.6	24.0	25.9	

*Significant at .10 level
**Significant at .05 level
***Significant at .01 level

Table T-5.4
Limited or Partial (t > 1) Cost Models, New York--1974, 1975 and 1976
(standard errors in parentheses)

Independent Variable	1974	1975	1976	Beta Coefficients 1975
		OUTPUT CHARACTERISTICS		
Scale				
BEDS1	x	x	-0.0553 (0.0480)	
BEDS2	x	x	x	
BEDS3	x	x	x	
BEDS4	x	x	x	
BEDS5	x	0.00689 (0.00536)	x	0.0311
Patient Age Distribution				
%UNDER65	x	x	x	
%AGE75 to 84	x	x	x	
%OVER84	x	x	x	
%FEMALE	0.0427* (0.0228)	x	0.0463 (0.0282)	
Patient Functional Status Variables (percentages of patients requiring assistance)				
%WALK	0.0387** (0.0152)	x	0.0280* (0.0157)	
%EAT		0.0428*** (0.0119)	x	0.122
%BATHE		0.0315* (0.0168)	0.0179 (0.0168)	0.0570
%BREATHE		-0.00757 (0.00618)	x	0.0308
%DRESS		x	0.0154 (0.0150)	
%CORRESPOND		0.0223* (0.0114)	x	0.0700
%MEDICATE		x		
%MASSAGE		x	x	

*Significant at .10 level.
**Significant at .05 level.
***Significant at .01 level.
x--Variables omitted that had been included in the "complete" models, Table T-5.3.

Table T-5.4 (cont'd.)

Independent Variable	1974	1975	1976	Beta Coefficients 1975
Patient Impairment Variables (percentages of patients with impairment)				
%ALERT			-0.0160 (0.0105)	
%SEE			0.0194* (0.0112)	
%HEAR			-0.0124 (0.0112)	
%SPEAK			x	
%COMMUNICATE			x	
%DEATHS	x		x	
ADMRATE	297* (172)	967* (130)	x	0.306
Service Offerings (0-1)				
DRUGS		x		
PHYS THERAPY		x		
OCC THERAPY		0.780 (0.506)		0.0369
SPEECH THERAPY		x		
SPECIAL DUTY		x		
PHYS SERVICES		0.986 (0.605)		0.0483
OTHER SERVICES		0.456*** (0.265)		0.0570
Service Intensity				
#PRESCRIPTIONS		x	0.0117* (0.00685)	
#X-RAYS		0.356*** (0.103)	0.592*** (0.0932)	0.0883
#PHYSICAL THERAPY		3.34*** (0.884)	1.30** (0.530)	0.115
#SPEECH THERAPY		x	12.4* (7.09)	
#OCC THERAPY		x	x	0.0478
#INHAL THERAPY		12.5** (6.61)	x	
#RADIATION THERAPY		x	35.8 (28.6)	0.102
#SOC SERVICEe		1.01*** (.376)	x	

*Significant at .10 level.
**Significant at .05 level.
***Significant at .01 level.
x--Variables omitted that had been included in the "complete" models, Table T-5.2

Table T-5.4 (cont'd.)

Independent Variable	1974	1975	1976	Beta Coefficients 1975
Bedroom Distribution				
%1BED		0.0300*** (.0149)	x	0.0679
%3BED		-0.0186* (.0109)	-0.0234*** (.0118)	0.0483
		INPUT PRICE VARIABLES		
LPN WAGE		0.687*** (.176)		0.104
RETAIL WAGE	6.25*** (1.76)	6.27*** (1.63)	7.78*** (1.87)	0.179
POPDEN	-.0000479 (.00003)	x	x	
		REGULATORY-TYPE VARIABLES		
HRF	-8.70*** (1.02)	-8.91*** (1.24)	-11.7*** (1.10)	-0.393
Sponsorship				
PROFIT	-9.25*** (0.966)	-11.6*** (1.29)	-10.2*** (1.05)	-0.556
NONPROFIT	-2.09** (1.04)	-5.51*** (1.09)	-5.87*** (1.16)	-0.244
Payor Distribution				
%PRIVATE	-0.0421*** (0.0131)	-0.0311** (0.0150)	-0.0653*** (0.0163	-0.0626
%MEDICARE	0.285*** (0.0800)	0.335*** (0.0915)	0.278*** (0.0902)	0.0938
Survey Certification Ratings				
NI	x	-2.16** (0.983)	-1.87* (0.968)	-0.0526
DEFERRED	x	x	x	
# UR			x	

*Significant at .10 level.
**Significant at .05 level.
***Significant at .01 level.
x--Variables omitted that had been included in the "complete" models, Table T-5.2.

Table T-5.4 (cont'd.)

Independent Variable	1974	1975	1976	Beta Coefficients 1975
Bed-Size Categories				
SIZE2	x	-0.7166 (0.531)	x	-0.0346
SIZE3	x	x	1.01* (0.559)	
SIZE4	x	x	x	
SIZE5	4.33*** (1.12)	x	2.40** (1.19)	
Regions				
REGION2	x	-0.911 (0.903)	1.66* (0.930)	-0.0259
REGION3	x	x	x	
REGION4	x	x	x	
REGION5	4.61*** (1.01)	3.49*** (0.997)	4.98***[a] (0.956)	0.116
REGION6	5.53*** (1.06)	6.03*** (1.14)		0.179
REGION7	6.94*** (1.17)	5.22*** (1.21)	5.24*** (1.24)	0.235
TECHNOLOGY OR PRODUCTION PROCESS VARIABLES				
SNF-HRF		-1.76** (0.656)	-0.906 (0.637)	-0.0734
HRF-SNF		-1.52* (1.06)	x	-0.0612
CLINIC		-2.66 (1.86)		-0.0326
%CONTRACT		-0.0633 (0.0483)		-0.0310
FACILITY AGE		-0.0195* (0.0118)		-0.0671
INDICATORS OF SHORT-RUN DISEQUILIBRIUM				
Occupancy Segments				
OCC1	x	x	x	
OCC2	-0.267*** (0.101)	x	-0.418*** (0.130)	
OCC3	x	x	x	
OCC4	0.234 (0.154)	x	-0.237* (0.132)	
CONSTANT	11.1	7.99	19.3	

*Significant at .10 level.
**Significant at .05 level.
***Significant at .01 level.
[a] Regions were combined for grouping purposes in 1976.
x--Variables omitted that had been included in the "complete" models, Table T-5.2.

Table T-5.4 (cont'd.)

Independent Variable	1974	1975	1976	Beta Coefficients 1975
Summary Statistics				
Coefficient of Determination (R^2)	0.644	0.764	0.758	
Corrected Coefficient of Determination	0.631	0.747	0.743	
F-Ratio	52.1	46.0	50.5	

reader should note that many other variables which do not satisfy the conservative 0.10 significance test are nevertheless significant at slightly lower levels-- e.g., at the 0.20 or 0.30 levels.

An examination of the results for individual variables and groups of variables reveals that many of the coefficient estimates are surprisingly similar across years despite the differences in model specification.

Recall that both continuous piece-wise variables (BEDS1-BEDS5) and categorical (SIZE2-SIZE5) bed-size dummy variables were defined--the first to explicitly investigate scale effects and the second to control for the use of bed-size as a grouping variable for rate-setting purposes (i.e., a regulatory variable). Although a variety of effects might conceivably be hypothesized for the grouping variables, one possibility is that costs gravitate toward the average cost level within the category and the "true" scale effects are distorted within each group. In any event, it was not clear a priori that scale and regulatory effects could be disentangled. Our efforts were exploratory in nature. It quickly became apparent that the multi-collinearity with respect to nine bed-size variables posed a serious problem. We dealt with this problem by entering all continuous (BEDS1-BEDS5) variables but entering only those categorical (SIZE2-SIZE5) variables having coefficients larger than their standard errors (t>1). Examining the results, we find that only one bed-size variable (SIZE5) is significant at the .10 level in 1974, none are significant in 1974, and two (BEDS1 and SIZE4) are significant in 1976. The estimated bed-size relationships for each year are portrayed graphically in Figure T-5.2. Statistical significance aside, there is no consistent pattern across years; it appears that the relationship is highly sensitive to differences in model specification.[1] Since the 1975 model is the most nearly complete, we conclude from the equation that economies or diseconomies of scale with respect to bed-size are minimally important cost factors, although the estimated coefficients indicate very modest economies of scale up to 100 beds and modest diseconomies thereafter. The 1976 model indicates the existence of substantial economies of scale over the lowest bed-size range (0-50 beds), a finding which is supported by the stepwise

---

[1] This conclusion was further supported from exploratory analysis that used a much less complete variable set and gave a more conventional U-shaped cost curve. Such results from earlier modeling efforts may be due to an incomplete specification of the output vector.

Table T-5.2: Bedsize Relationships Estimated from the Complete Model --1974, 1975 and 1976

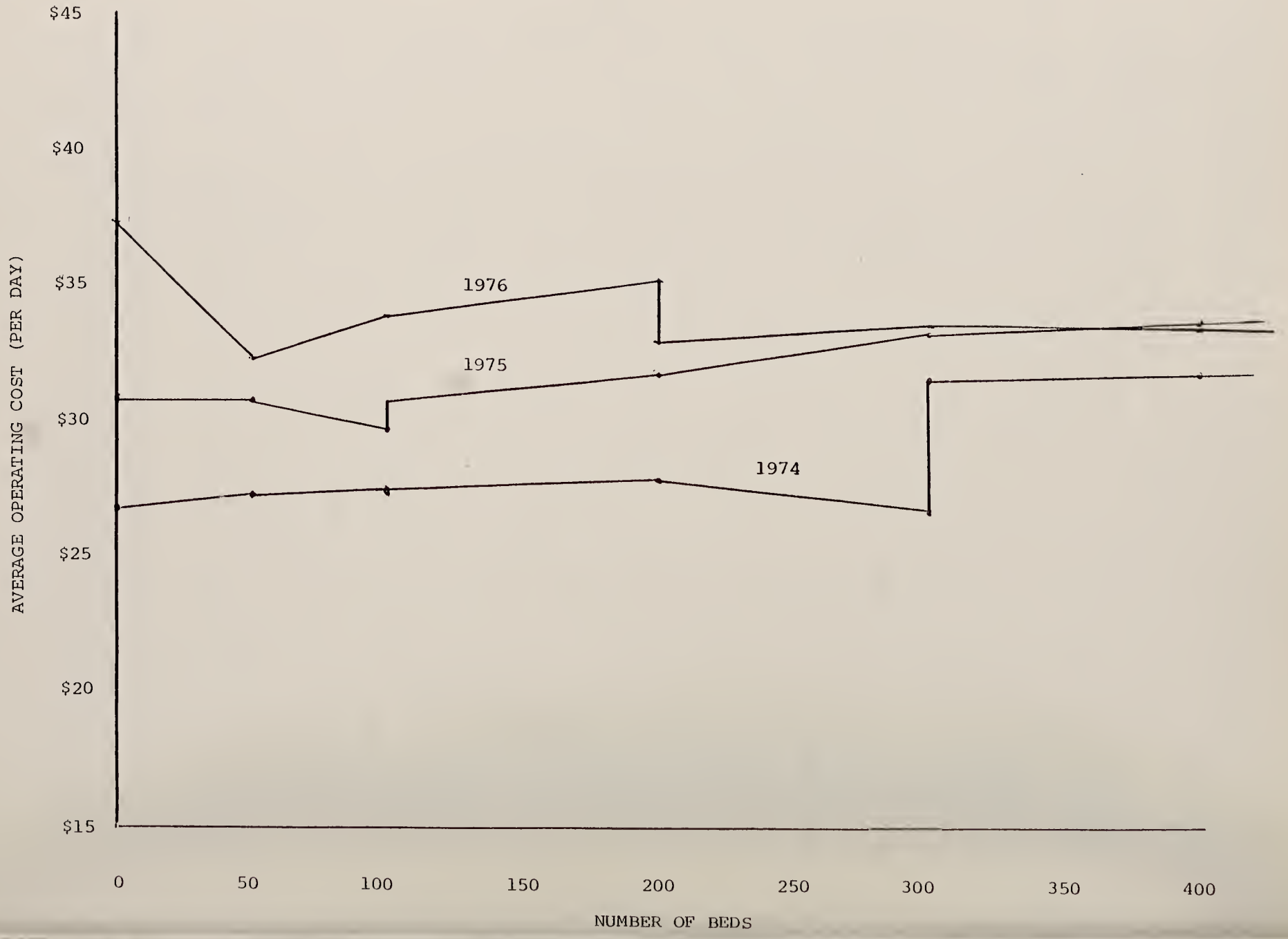
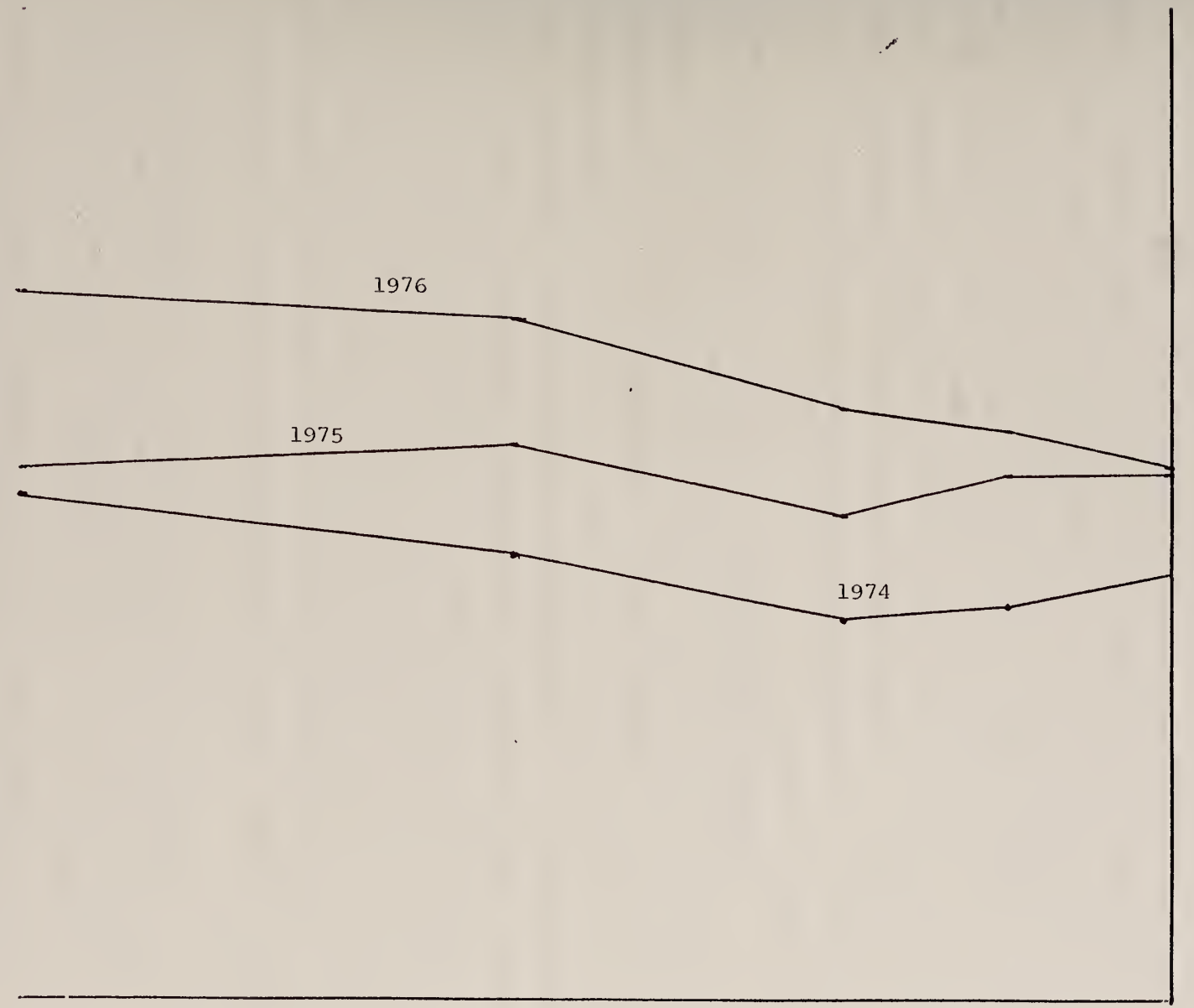


Table T-5.3: Occupancy Relationships Estimated from the Complete Model --1974, 1975 and 1976



(t>1) modeling efforts (see the next section). None of the conclusions with respect to bed-size were substantially changed when categorical bed-size variables were excluded--i.e., when only BEDS1-BEDS5 were used.

Occupancy. The results with respect to the occupancy variables (OCC1-OCC4) are somewhat more consistent with the usual expectations. Although none of the variables are ever significant, all three models suggest slightly decreasing costs between 80 and 90 percent occupancy as indicated by a negative coefficient for the OCC2 variable. Other results are not so uniform across the three years. The 1974 and 1976 models indicate decreasing costs for less than 80 percent occupancy while the 1975 model indicates very nominal increasing costs. Only the 1976 model indicates decreasing costs above 90 percent occupancy whereas the 1974 and 1975 models indicate increasing costs. (It is not uncommon to find increasing costs as production nears capacity.) In any event the results with respect to occupancy were neither dramatic nor significant as shown in Table T-5.3, possibly because the capital and other more nearly fixed costs were not included in the AOC dependent variable.

Regulatory-Type Variables. In general, the coefficients of the regulatory-type variables are strikingly similar and consistently significant from year to year. After controlling for differences in patient and service mix, average per diem operating cost is nine to eleven dollars less in a Health Related Facility (HRF) than it is in a Skilled Nursing Facility (SNF). For-profit facilities are estimated to cost nine to twelve dollars less than otherwise similar public facilities; non-profit facilities are estimated to cost six dollars less than public facilities in the 1975 and 1976 models, although the less complete 1974 model indicates that the difference is only about two-and-one-half dollars. All of these results are highly significant. It is interesting that the estimated cost differential between non-profit and for-profit homes declined from seven dollars in 1974 and six dollars in 1976, the year in which sponsorship ceased to be a grouping variable.

Although the results for the survey certification variables--NI ("needs improvement") and DEFERRED--are consistent across years, only the NI variable for 1975 and 1976 is significant.[1] The 1975 and 1976 results

[1]It is not surprising that NI is not significant in 1974 for the survey ratings actually related to the situation at the beginning of 1976.

indicate that NI facilities, facilities needing improvement, cost about $2.00 to $2.50 less, presumably due to not having incurred the costs of satisfying the certification requirements. The DEFERRED or rating "deferred" facilities were estimated to cost about $1.00 more in 1976, although the 1974 and 1975 models indicated smaller cost increments.

The models consistently estimate that private patients cost less than otherwise similar Medicaid patients and that Medicare patients cost significantly more than either private or Medicaid patients. As an illustration, on the 1975 results, the estimates imply that the per diem cost would be $4.06 less for a facility having 100 percent private patients, compared to a facility having all Medicaid patients, and that the per diem operating cost would be almost $35 higher for a facility having all Medicare patients. Inasmuch as we have already controlled for patient characteristics, service mix, and certification difference, it is not clear what differences in care, if any, result from the increased cost of public patients. It is also important to note that the result with respect to private patients was not found from analysis of the national data in Technical Paper #4. Apparently the Medicaid program in New York is more "generous" than it is in most other states. This finding suggests that a meaningful reimbursement formula cannot be derived from a nationwide analysis--i.e., the programmatic differences between states may be too large to be reasonably captured in a nationwide modeling effort.

Controlling for regional wage levels--and in 1975 for facility wage levels--there are consistently significant coefficients for several of the location dummy variables (REGION1-REGION7). In particular, compared to the West region, the North Metro (REGION5), Long Island (REGION6) and New York City (REGION7) regions cost considerably more, as much as six dollars more per day. Again, it is not clear what differences in care correspond to these cost differences since patient and service mix differences have been controlled for. Finally, there is no statistically significant relationship between costs and the intensity of utilization review (#UR).

Patient Variables. The patient-age distribution variables (%UNDER65, %AGE75TO84, and %OVER84) are insignificant in all three years; functional status and impairment variables apparently reflect relevant age-specific differences. Percent female (%FEMALE) is positively related to cost--as was found by the AMS study--although the coefficient is insignificant in any year.

The 1974 model included one functional status variable (%WALK). The 1975 model included eight functional status variables (%WALK, %EAT, %BREATHE, %DRESS, %CORRESPOND, %MEDICATE, AND %MASSAGE). The 1976 model included seven (excluding only %MEDICATE). These variables indicated the percentages of patients requiring either partial or full assistance and we predicted positive signs for all of them. (See Table T-5.2 for definitions). The one variable used in 1974 is positive and significant. In 1975 four of the eight variables are positive and the three significant variables (%EAT, %BREATHE, and %CORRESPOND) are all positive. In 1976 five of the seven ambulatory status variables are positive, although only the positive coefficient on %WALK even approaches significance. These results may reflect substantial multi-collinearity. The partial models reported in Section T-5.4 are less likely to suffer this problem.

For 1976 only, there are five patient impairment variables (%ALERT, %SEE, %HEAR, %SPEAK, and %COMMUNICATE)--the percentage of patients with either partial or complete impairment in each. Only the positive coefficient on %SEE is significant at the 0.10 level, although the negative coefficient on %ALERT is significant at better than the 0.20 level. This latter result is not anomalous. It is not obvious that patients with impairment should cost more; such patients, especially those not so alert, might be less demanding or more easily ignored by the staff.

In both 1974 and 1976 percentage deaths (%DEATH) was found to be negatively but insignificantly related to cost. The variable was not useable in 1975.

The cost per admission--the coefficient for ADMRATE variable--is estimated to be $342 in 1974, $276 in 1976, and $1,050 in tne more fully-specified 1975 model. Only the 1976 coefficient is insignificant.

Quality Proxies. In both 1975 and 1976, but not 1974, there is information on seven service intensity variables (#PRESCRIPTIONS, #X-RAYS, #PHYS THERAPY, #SPEECH THERAPY, #OCC THERAPY, #INHAL THERAPY, #RADIATION THERAPY and #SOC SERVICES) which might be thought to reflect differences in the "quality" of care. Of course, homes providing greater serviced intensity will (in theory) have higher costs, although multi-collinearity might be a problem here as well. In 1975, seven of the eight coefficient estimates

are positive as predicted. Furthermore, of the four that are significant (#X-RAYS, #PHYS THERAPY, #INHAL THERAPY and #SOC SERVICE), all are positive. In 1976, seven of the eight coefficient estimates are again positive. Four variables (#X-RAYS, #PHYSICAL THERAPY, #PRESCRIPTIONS and #SPEECH THERAPY) are again significant and positive, although only two of the variables are the same as those in 1975. In 1975 only, we had additional information on the offering of services and their inclusion in the Medicaid rate (DRUGS, PHYS THERAPY, OCC THERAPY, SPEECH PATH, SPECIAL DUTY, PHYS SERVICES, OTHER SERVICES). (See Section T.5.2 for discussion and motivation of these variables.) Only the coefficient obtained for OTHER SERVICES was significant and positive as expected. However, positive coefficients for OCC THERAPY and PHYS SERVICES are also significant at a level slightly lower than the usual 0.10 level.

The percentages of patients in either one or three (or more) bed rooms--%1BED and %3BED--were included as quality measures of a less clinical nature. The 1975 results were the most promising. As expected, one-bed rooms cost more and three-bed rooms cost less when compared to two-bed rooms. The results for 1976 are not significant and the data were not available for 1974.

Price Variables. The average retail wage (RETAIL WAGE) in the county of facility location is used as an index of regional wage levels. The resulting coefficient estimates are uniformly positive, as expected, and highly significant. In fact, the estimated coefficients are virtually identical. In 1975 there was information sufficient to calculate the average LPN wage (LPN WAGE) within each facility. This variable also gave a significant and positive coefficient although it is about one-tenth the size of the coefficient for RETAIL WAGE.[1] The county-wide population density (POPDEN) was further included as a control for systematic price differences in more urban or populated areas. The estimated coefficient is insignificant in the more fully-specified 1975 and 1976 models.

---

[1]The multi-equation modeling found evidence that nursing wage levels within facilities were endogenous, partly reflecting skill-level differences related to variation in service-mix.

Technology Variables. In 1975 average operating cost is negatively related to facility age (FACILITY AGE). Although the coefficient is insignificant, this result is somewhat unexpected. It was expected that newer facilities would be more efficient to operate and less costly to maintain. On the other hand, newer facilities may also support a higher amenity level of care--e.g., air conditioning or individually-controlled thermostats. The 1975 model yields negative but insignificant coefficients with respect to the percentage of services contracted (%CONTRACT) and for facilities maintaining an independent clinic (CLINIC). Initial hypotheses were partially supported. Facilities appear to contract for those services that can be obtained more cheaply elsewhere and economies accrue to those facilities having an independent clinic attached. Finally, both 1975 and 1976 models indicate that combined facilities (SNF-HRF and HRF-SNF) cost less, although the result is only significant in 1975 for SNF-parts of combined facilities (i.e., the SNF-HRF variable). Nevertheless, these findings are consistent with the premise that economies accrue to joint SNF and HRF production. If costs were merely shifting between the SNF- and HRF-parts, opposite signs would be obtained for the two variables.

Groups of Variables. Whereas the models discussed here give a "good fit", the interpretation of the coefficients is somewhat uncertain given the degree of multi-collinearity resulting from the large number of independent variables--many of them measuring related factors. In the following paragraphs we report on attempts to assess the significance of groups of variables. The tests employed are not as sensitive to multi-collinearity problems as the individual t-tests used above.

The coefficient estimates obtained for certain variables may partly reflect the effects of other (correlated) variables, either omitted variables or variables actually included in the model. In our analysis entire groups of variables were tested for statistical significance using a residual F-test. The results for 1975 are reported in Table T-5.5. The variable groups are significant at the .05 level or better if the F-statistic is greater than or equal to the indicated criterion value. The following variable groups are significant at the .05 level: (1) ambulatory status variables, (2) service intensity variables, (3) bedroom distribution variables, (4) price variables, (5) sponsorship variables, (6) certification

Table T-5.5
RESIDUAL F-TEST RESULTS

Variable Group	F-Statistic	Significance Criterion for .05 Level
Bed-Size (BEDS1 -BEDS5 and SIZE 2-SIZE %)	0.94	1.85
Occupancy (OCC1-OCC4)	0.54	2.23
Bed-Size and Occupancy	0.74	2.23
Ambulatory Status (%WALK, %EAT, %BATHE, %BREATHE, %DRESS, %CORRESPOND, %MEDICATE and % MASSAGE)	4.44	1.96
Age and Sex Distribution (%UNDER65, %AGE75TO84, %OVER84 and %FEMALE)	0.05	2.30
Service Intensity (#PRESCRIPTION, #X-RAYS, #PHYS THERAPY, #OCC THERAPY, #INHAL THERAPY, #RAD THERAPY, and #SOC SERVICES)	5.39	1.96
Service Offerings (DRUGS, PHYS THERAPY, OCC THERAPY, SPEECH THERAPY, SPECIAL DUTY, PHYS SERVICES and OTHER SERIVCES)	1.35	2.03
Bedroom Distribution (%1BED and %3BED)	3.98	3.02
Input Prices (LPN WAGE, RETAIL WAGE, and POPDEN)	8.54	2.62
Sponsorship (PROFIT and NONPROFIT)	33.65	3.02
Certification (HRF Only)	36.00	3.86
Payor Distribution (%PRIVATE and %MEDICARE)	7.55	3.02
Survey Certification Rating (NI and DEFERRED)	2.68	3.02
Technology or Production Process (CLINIC, %CONTRACT and FACILITY AGE)	2.67	2.23
Region (REGION2 through REGION7)	5.74	2.23

variables, (7) payor distribution variables, (8) region dummies, and (9) production process variables. The age and sex variables are insignificant. The bed-size, occupany, and bed-size and occupancy combined categories are also insignificant. The service offering variables are also insignificant. Finally, although the survey certification variables are not significant at the .05 level, they are significant at the less conservative .10 level. However, the reader should bear in mind that these results may depend substantially upon the specifications of the model within which the variable groups have been tested. If, for example, the least significant variables groups (age and sex, bed-size and occupancy) were omitted, a residual F-test might well reveal that the service offering variables were significant.[1]

In an effort to diagnose the quantitative importance of entire groups of variables, the absolute values of the beta coefficients have been summed to give an ad hoc index of the extent to which different kinds of variables explain observed variation in the average operating cost levels.[2] (In a sense, the beta coefficients scale the results for different variables to a common denominator.) The summed absolute values for 1975 are given in Table T-5.6 by variable group; the individual beta coefficients can be found in Table T-5.3.

---

[1] A residual F-test was also conducted on the patient impairment variables in the 1976 model. These variables were found not to be significant ($F - 1.04$, $CV_{.05} = 2.23$).

[2] The beta coefficients are simply normalized regression coefficients and bear a straightforward relationship to the unnormalized regression coefficients, being equal to the unnormalized coefficient times the ratio of standard deviations for the independent (S_x) and dependent (S_y) variables:

$$b_x \cdot \frac{S_x}{S_y}$$

A beta coefficient of, say, .25 can be interpreted to mean that a one-standard deviation change in the independent variable leads to a .25 standard deviation change in average operating costs, the dependent variable. (It is interesting to note that the beta coefficient for the independent variable in a bivariate regression model is identically equal to the simple correlation between the two variables.)

Given that a higher beta sum implies that a variable set is more important,[1] three groups of regulatory-type variables--sponsorship, certification, and region--have the greatest quantitative importance. (All of these are or were official New York grouping variables.) Other important variable groups are bed-size and occupancy combined, ambulatory status, service intensity and input prices. Less important are the service offerings, bed room distribution, payor distribution, technology, and bed-size or occupancy taken separately. The age and sex distribution and survey certification variables are the least important.[2] Of course, these also may be highly sensitive to the particular variables being used. For example, the "quality of care"--aside from service offerings and service intensity--may be a highly important variable, albeit one that is much too crudely measured from the survey certification ratings. Furthermore, the groupings of variables may not be the most appropriate ones. If, for example, the service intensity and service offering variables are combined, services then become the second most important set of variables. (The reader may wish to combine other variable groups.) Perhaps the most noteworthy finding from this ad hoc analysis is the indication that bed-size and occupancy are quantitatively important even though a residual F-test indicates that the variables are not significant as a group. Because of this result we are inclined to discount the significance test. As with other groups of variables, the dimensionality upon which the F-test is based is somewhat artificial. Although a total of six bed-size variables were used, a lesser number of bed-size variables might well yield a significant relationship. We explore such prospects in the next section.

---

[1] One drawback of this approach to assessing the importance of groups of variables is that it overstates variable importance when multi-collinearity is present in the group.

[2] The sum of beta coefficients for the patient impairment variables in the 1976 model is 0.11.

Table T-5.6

Summed Beta Coefficients

Variable Group	Beta Coefficient Sum
Bed-Size	0.19
Occupancy	0.18
Bed-size and Occupancy	0.37
Ambulatory Status	0.34
Age and Sex Distribution	0.38
Service Offerings	0.28
Bedroom Distribution	0.11
Price	0.30
Sponsorship	0.80
Certification	0.39
Payor Distribution	0.16
Survey Certification Ratings	0.07
Region	0.52
Technology or Production Process	0.24

T-5.3.2 The Limited or Partial Models

This section addresses the general multi-collinearity concern by reporting models that include only those variables entering the relationship with a t-statistic larger than one--i.e., having a coefficient larger than its standard error. Such selection of the independent variables is by no means "ideal," however, within the scope of this contract there was no satisfactory alternatives for reducing the dimensionality of the model and thereby diminishing the degree of multi-collinearity.[1] Perhaps future studies can resolve the dimensionality problem more satisfactorily by developing and indicating composite indices for patient- and service-mix costliness, instead of using the many variables needed to describe the actual differences in patient and service-mix. However, such "costliness" indices were not available and quite likely cannot be determined from statistical analysis alone.

[1] Although some econometricians are reluctant to approve such empirically-based strategies for model specification, the following from Wonnacott and Wonnacott is noteworthy:

> ". . . if there are clear prior guidelines indicating that a few specific regressors are appropriate, then they should be used right away in a full multiple regression, rather than tested one at a time with any sort of stepwise approach. If there are no such prior guidelines, but the number of regressors must be kept small to provide a more manageable model, then a stepwise technique may be reasonable." (p. 312)

The limited or partial (t>1) models for 1974, 1975 and 1976 are shown in Table T-5.4. These models are exceptionally robust vis-a-vis the more complete or "fully-specified" models given in Table T-5.2.[1] Interestingly, the percentages of variation explained in the "limited" models are virtually the same as those for the corresponding "complete" models: 64.4 percent versus 64.7 percent in 1974, 76.4 percent versus 76.8 in 1975 and 75.8 percent versus 76.4 percent in 1976. The corrected coefficients of determination and F-ratios are actually higher in the limited models. Nevertheless, the limited and complete models could have differential implications for reimbursement. For example, the certification and sponsorship variables in a limited model might partly reflect the patient and service variables omitted for lack of statistical significance. On the other hand, if the certification, sponsorship and other coefficients do not change, a reasonable conclusion is that such "spurious" correlation is not a serious problem for rate setting and that the effect of omitted variables is reflected by the more relevant (non-regulatory) patient and service-mix variables still included.

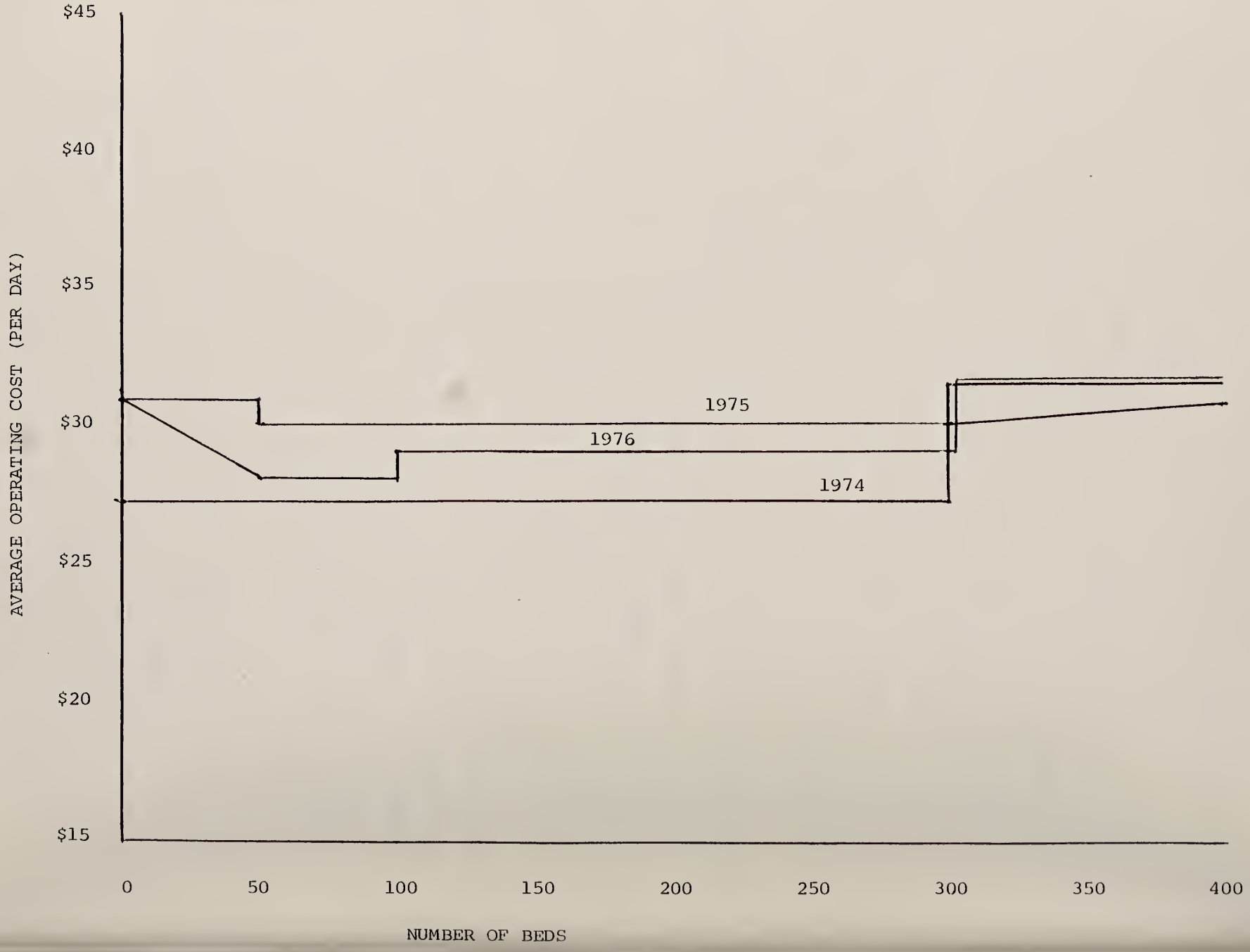
As expected, a large percentage of variables in these limited (t>1) models are significant at the 0.10 level or better. In 1974, 14 of the 16 included variables (or 88 percent) are significant. In 1975, 24 of the 33 included variables (or 73 percent) are significant. In 1976, 21 of the 29 included variables (or 72 percent) are significant. The number of significant variables are about the same as those in the complete models--somewhat higher in 1974 and 1975, and lower in 1976. Examined below are the results for individual variables and groups of variables.

Bed-Size. No continuous bed-size variables were included (i.e., entered with a coefficient larger than its standard error) in the 1974 model and one each were included in the 1975 and 1976 models--BEDS5 and BEDS1, respectively. However, neither coefficient is significant. Categorical bed-size variables were included in each of the three years--SIZE5 in 1974, SIZE2 in 1975, and SIZE3# and SIZE5 in 1976. Only the 1974 and 1976 results are significant. The estimated bed-size relationships are portrayed graphically in Figure T-5.4. A more coherent pattern emerges from these results

---

[1]In the formal sense, any more limited model, however obtained, should be regarded as giving a more partial model and we claim no more than that for the models presented here.

Table T-5.4: Bedsize Relationships Estimated from the Stepwise (t>1) Model--1974, 1975 and 1976



than did from the results for the "complete" model. All three years indicate decreasing returns to scale beyond 300 beds; all three indicate constant returns to scale between 100 and 200 beds; and the two more fully-specified models (1975 and 1976) indicate increasing returns to scale between 0 and 50 beds. The 1976 model also indicates a significant cost increment of one dollar per day above 100 beds. It is also noteworthy that these results are more nearly consistent with the hypothesis that grouping by bed-size has distorted the scale effect inasmuch as a larger number of categorical than continuous variables entered the relationships.

Occupancy. Considered here are findings for the occupancy segments (OCC1-OCC4), the proxies for short-run disequilibrium. Two of these variables (OCC2 and OCC4) were included in the 1974 model; none were included in the 1975 model; and two (OCC2 and OCC4, again) were included in the 1976 model. The estimated occupancy relationships are portrayed graphically in Figure T-5.5. As found from the complete models, both the 1974 and 1976 limited models indicate decreasing costs between 80 and 90 percent occupancy. However, also as found from the complete models, the 1974 limited model indicates increasing costs between 95 and 100 percent occupancy whereas the 1976 model indicates decreasing cost. In any event, the 1975 finding of constant returns to occupancy are most appropriate since facilities changing bed-size over the year were explicitly identified in that year and removed from the data. Inability to do this in 1974 and 1976 may well have biased our estimates inasmuch as the occupancy levels were calculated as average daily census divided by bed-size. If, for example, bed-size had increased over the year the occupancy figure obtained would be biased downwards and might well proxy the costs associated with a change in bed-size.

Regulatory-Type Variables. The findings with respect to other regulatory variables are strikingly similar to those obtained from the complete models. To begin, the HRF, PROFIT, NONPROFIT, %PRIVATE and MEDICARE variables entered all three models at highly significant confidence levels and the coefficient estimates are virtually the same as those from the complete models. For example, the HRF coefficient for 1975 is -8.82 in the limited model and -8.91 in the complete model; the PROFIT coefficient is -11.6 in the limited model and -11.6 in the complete model; and the %PRIVATE coefficient is -.0364 in the limited model and -.0311 in the complete model.

Table T-5.5: Occupancy Relationships Estimated from the Stepwise (t>1) Model--1974, 1975 and 1976

The DEFERRED survey certification variable never enters the limited models. Its coefficient in the complete model had likewise never been larger than the standard error of estimate. However, the NI ("need improvement") variable entered significantly in both 1975 and 1976, the same years in which it had been significant in the complete models. Again, the coefficient estimates are very similar. The utilization review intensity variable (#UR) did not enter in 1976, and the only year for which it was available. It was insignificant in the complete model.

Finally, the results with respect to the regional grouping variables (REGION2-REGION7) are also similar. Only variables obtaining coefficients larger than their standard errors in the complete model entered the limited models, and the same variables are significant. Although coefficient estimates vary by as much as a dollar in several cases (e.g., REGION7 in 1975) the differences are not sizeable.

Patient Variables. Patient age variables (%UNDER65, %AGE75TO84 and %OVER84) were not included in any of the limited models. They also had been highly insignificant in the complete models. Except for the highly insignificant variables, which were omitted in the limited models, the results for the functional status variables are essentially unchanged from those obtained in the parallel complete models. Signs, significance levels and the actual coefficient estimates remain approximately the same. The one functional status variable (%WALK) in 1974 is again positive, as predicted, and significant. Three (%BATHE, %EAT and %CORRESPOND) of the four variables in the 1975 model are significant and positive. The fourth (%BREATHE) is negative, albeit not significant. Only (%WALK) of the three functional status variables in the 1976 model is positive and significant. However, the other two variables (%BATHE and %DRESS) are also positive as hypothesized. Recall that the patient impairment variables were only available for 1976. Once again, the limited model results are not meaningfully different from results in the more fully-specified model. Only %SEE entered significantly, as it had in the complete model. The parameter estimate is 0.0194 compared to 0.0204 from the complete model. Both %ALERT and %HEAR entered with negative but insignificant coefficients, as they had in the complete model.

The %DEATHS variable did not enter either the 1974 or 1975 models for which this variable was available; it had also been insignificant in the

complete models. The results obtained for the ADMRATE variable in 1974 and 1975 are very similar to those obtained for those years in the complete models. It did not enter in 1976; the coefficient estimate from the complete model for that year had barely exceeded its standard error of estimate.

"Quality" Proxies. Four (#X-RAYS, #PHYS THERAPY, #OCC THERAPY and #SOC SERVICES) of the seven service intensity variables entered the 1975 model and five (#PRESCRIPTIONS, #X-RAYS, #PHYS THERAPY, #SPEECH THERAPY, and #INHAL THERAPY) entered in 1976. All nine coefficient estimates were positive, as predicted, and only #INHAL THERAPY in 1976 was not significant at the 0.10 level. For all variables, the limited equation parameter estimates were similar to those from the so-called complete equation. We found much the same with respect to the service offering variables in 1975. The results for both the complete and limited models are nearly identical. The SOC SERVICES variable is significant but the OCC THERAPY and PHYS SERVICES variables are insignificant. All coefficients are positive, as expected. Finally, except for %3BED in 1975 which became significant at the 0.10 level, the results for the bedroom distribution variables are also similar to those from the complete model.

Price Variables. Here also, the limited model results are very much like those from the complete model. The population density (POPDEN) variable is omitted in 1975 and 1976; however, it had also been highly insignificant in the complete models for those years.

Technology. Except for the omission of HRF-SNF in 1976, which had also been insignificant in the complete model that year, the results are once again virtually indistinguishable from those for the complete models.

Summary. The pattern of similarity between coefficient estimates and significance obtained in the limited and complete models leads inescapably to the conclusion that multi-collinearity has not been so serious a problem in the complete models as might have been thought given the large number of variables included. Moreover, since the limited models give virtually the same "fit" to the data, it also suggests that nothing is gained empirically by using the more complete models. The effect of omitting variables in the limited models seems to have been little more than to trivially increase the standard error of estimate.

T-5.4 Pooled Cross-Section/Time Series Analysis--A First-Difference Model

This section reports the results of estimating pooled cross-section/time series models for the first difference or change in average operating costs between 1974 and 1976. In general, the first-difference model purges the data of so-called "steady state" effects--i.e., factors peculiar to a facility that persist over time and which may be correlated with other characteristics of the facility. The first difference model also allows inclusion of reimbursement-related variable, the first-difference in the actual reimbursement rate, without risking serious bias due to the "steady state" effects subsumed in that variable. Although the reimbursement rate would have been a conceptually appropriate variable to include in the cost function, it was omitted for the following reason. Except for facilities subject to ceiling limitations, the rate was set equal to average total cost--both operating and capital expenses--times some constant inflation factor. Thus, the variable would substantially reflect uncontrolled facility-specific differences that might otherwise enter the disturbance term (i.e., state effects) and we might wrongly construe the results as reflecting reimbursement policy. However, as noted above the first-difference specification does not suffer this problem. It does have another less serious problem--namely, that the rates used in computing the first difference rate reflect total cost whereas concern throughout is with only the operating cost component. This limitation reduces the consistency of estimation. It also introduces a source of bias.[1]

Although we are not explicitly estimating a cost function, the underlying theoretical cost function determinants are unchanged for those indicated in Section T-5.1. However, since only a limited number of variables were available for all three years, the operational model is considerably less ambitious. It relates the year-to-year percentage change in average operating cost (ΔAOC) to the following list of change variables:

---

[1] The first difference reimbursement rate is measured with error as a result of the way it was computed. Least squares estimates of the regression parameters will be biased and inconsistent with the degree of bias and inconsistency being related to the variance of the measurement error. See Johnson (1972), p. 281-291.

ΔRATE--percentage change in the reimbursement rate;

ΔOCC--change in percentage occupancy;

ΔMEDICAID--change in percentage Medicaid patients;

ΔPRIVATE--change in percentage private patients;

ΔAGE--change in mean patient age;

ΔADMRATE--percentage change in the admission rate;

ΔFEMALE--change in percentage female; and

ΔWALK--change in percentage nonambulatory patients.

A new regulatory variable (OVER), also was defined as the percentage by which average total cost in the preceding year exceeded the Medicaid rate. As such, this variable indicates the extent, if any, to which a facility is "bumping up" against ceiling limitations in rate setting. A fundamental precept of prospective reimbursement is that facilities not being fully reimbursed will moderate their rate of cost increase. Included variables indicate the levels of all first-differences in the variables, except RATE, as well as the levels of other variables for which change data were not available (e.g., RETAIL WAGE). These variables were defined in Section T-5.2 and variable definitions are summarized in Table T-5.1. This analysis, in contrast to earlier ones uses the actual bed-size (BEDS) and occupancy (OCC) instead of "grafting" segments as done in Section T-5.3. Also, the average age (AGE) is used instead of the set of age variables. Since the sample size is reduced to only 225 cases after merging the three years of data, only those "absolute level" variables entering with a coefficient greater than its standard error ($t>1$) were included, thereby conserving on degrees of freedom and mitigating potential multi-collinearity problems.

The results from both 1974-1975, 1975-1976 and 1974-1976 analyses are shown in Table T-5.6. Overall, the three equations were significant at the 0.10 level and gave coefficients of determination ranging between 36.2 and 43.2 percent. The 1975-1976 model gave the highest F-ratio and also the highest corrected coefficient of determination. In the 1974-1975 model, 17 of the 40 entering variables (or 42 percent) were significant at the 0.10 level or better; in the 1975-1976 model 15 of the 35 entering variables (or 43 percent) were significant at the 0.10 level or better; and in the 1974-1976 model 24 of the 49 entering variables (or 49 percent) were significant at the 0.10 level or better. (If all variables were included in each of these models these ratios would be reduced by as much as half.)

Discussed first are the results for the ΔRATE and OVER variables, which are both explicit indicators of reimbursement policy. The ΔRATE variable is a percentage change in the Medicaid reimbursement rate. If prospective reimbursement is effective as a cost control mechanism, then cost should respond to exogenously-determined changes in the rate. The findings support that proposition. The 1974-1975 coefficient estimate is positive and significant at the 0.10 level, and the 1975-1976 coefficient estimate is positive and significant at a level slightly lower than the 0.10 level.[1] Although ΔRATE obtains a positive coefficient in the composite 1974-1976 model, it is insignificant. Results for this model might be improved if separate ΔRATE variables had been indicated for each of the two years.

Recall that OVER--actually OVER74 and OVER75--was defined as the percentage, if any, by which average total cost exceeds the Medicaid rate in the preceding year--either 1974 or 1975. This variable is a proxy for the extent to which a facility is "bumping up" against ceiling limitations in rate setting. If such ceilings are effective in moderating the rate of cost increase this variable should have a negative coefficient. A negative and significant coefficient is obtained in the 1974-1975 model. However, in the 1975-1976 model, there is a similarly significant positive coefficient. Although such dramatic reversal might ordinarily be regarded as evidence of model misspecification, another explanation is possible. Recall from Section T-5.1 that a new provision for waiver of cost ceilings was adopted in 1976. Therefore, it could be predicted that the efficacy of cost ceilings was weakened in that year. The result for the 1975-1976 model may suggest that the anticipation of such waivers has altogether thwarted the efficacy of cost ceilings, at least temporarily. Of course, its longer-run impact would depend upon the willingness of the regulatory authority to actually grant the waivers. Neither of the OVER variables is significant in the composite 1974-1976 model.

Consider the other first-difference variables in Table T-5.6. None of them are consistently significant in all three models. The ΔOCC variable

---

[1] As noted in Appendix T-5-A, facilities were reimbursed at their 1975 rate for at least the first ten months of 1976. It is not clear whether our data reflect reimbursement at the 1975 or 1976 rates. The data indicates an overall reduction in the rate between 1975 and 1976, suggesting that it reflects the expected 1976 rate. However, industry sources claim that this is unlikely and that most facilities would have reported actual revenues. We have not found anyone at the New York State Department of Health who could resolve this quandary.

Table T-5.6

FIRST DIFFERENCE COST MODELS, NEW YORK--1974-1975, 1975-1976 and 1974-1976

(standard errors in parentheses)

Independent Variable	1974-1975	1975-1976	1976-1977
		FIRST-DIFFERENCE VARIABLES	
ΔRATE	0.000729** (0.00034)	0.000182* (0.00011)	0.0000988 (0.00014)
ΔOCC		-0.000890 (0.00068)	0.00126 (0.00106)
ΔMEDICARE	0.00184 (0.00299)	0.00176 (0.00311)	0.747* (0.338)
ΔPRIVATE	0-.000748 (0.00118)	-0.000792 (0.000938)	-0.112 (0.103)
ΔAGE	0.00192 (0.00395)	-0.0128* (0.00357)	0.00533 (0.00493)
ΔADMRATE	0.000353* (0.00011)	-0.000176 (0.00012)	0.000373* (0.00020)
ΔFEMALE	0.000334 (0.00080)	0.00322* (0.00099)	0.00150 (0.00127)
ΔWALK	0.000442 (0.000342)	-0.000514 (0.000405)	-0.0000721 (0.000444)
OVER74	-0.000128* (0.00007)		-0.0000220 (0.000150)
OVER75		0.000137* (0.00005)	-0.0000880 (0.00010)
		OUTPUT CHARACTERISTICS	
BEDS	0.0000580 (0.00006)		-0.0000220 (0.000150)
AGE		-0.00296 (0.00216)	-0.0000880 (0.00010)
%FEMALE			
Patient Functional Status Variables (percentages of patients requiring assistance)			
%WALK		-0.000839** (0.000434)	
%EAT	0.000585* (0.00030)		0.000294 (0.00039)
%BATHE			
%BREATHE	-0.000172 (0.00013)		
%DRESS	-0.000740 (0.00045)		-0.0132** (0.00061)
%CORRESPOND	-0.00510* (0.00029)		
%MEDICATE	-0.00115* (0.00049)		-0.00120* (0.00072)
%MASSAGE		0.000242 (0.00024)	

*Significant at .10 level.

**Significant at .05 level.

***Significant at .01 level.

Table T-5.6 (cont'd.)

Independent Variable	1974-1975	1975-1976	1976-1977
Patient Impairment Variables (percentages of patients with impairment)			
%ALERT	-0.000532* (0.00029)		-0.000979** (0.00042)
%SEE			0.000969*** (0.00034)
%HEAR	0.000272 (0.00020)		
%SPEAK			-0.000831* (0.00042)
%COMMUNICATE			0.000546 (0.00040)
%DEATHS		-0.000351* (0.00021)	-0.00150*** (0.00052)
ADMRATE		-8.66** (4.11)	
Service Offerings (0-1)			
DRUGS		-0.0269*** (0.0103)	-0.0375*** (0.0173)
PHYS THERAPY			
OCC THERAPY	-0.0124* (.0117)		-0.0212 (0.0168)
SPEECH THERAPY			
SPECIAL DUTY		-0.117*** (0.0255)	-0.127*** (0.0384)
PHYS SERVICES	-0.0174 (0.0135)		
OTHER SERVICES	-0.00632 (0.00633)		-0.0296* (0.00886)
Service Intensity			
#X-RAYS	0.00202 (0.00213)	-0.00351* (0.00196)	-0.0356 (0.00304)
#PHYSICAL THERAPY	-0.0271 (0.0213)		-0.0361 (0.0294)
#SPEECH THERAPY		0.127 (0.0982)	
#OCC THERAPY			0.0814 (0.0528)
#INHAL THERAPY	-0.339* (.172)		
#RADIATION THERAPY			
#SOC SERVICE	-0.0151 (0.0131)		
Bedroom Distribution			
%1BED	0.000518 (0.00041)		0.00117** (0.00056)
%3BED		0.000309 (0.00024)	0.000913** (0.00037)

*Significant at .10 level.
**Significant at .05 level.
***Significant at .01 level.

Table T-5.6 (cont'd.)

Independent Variable	1974-1975	1975-1976	1976-1977
		INPUT PRICE VARIABLES	
LPN WAGE	0.00624 (0.00411)	-0.01067*** (0.00367)	
RETAIL WAGE	0.0263 (0.0326)	-0.0288 (0.0260)	
POPDEN	-0.00000242*** (0.000000)	0.0000013*** (0.000000)	-0.000000714 (0.000000)
		REGULATORY-TYPE VARIABLES	
HRF	-0.134*** (0.0414)	0.0569** (0.0282)	-0.140*** (0.0540)
Sponsorship			
PROFIT	-0.0944*** (0.0324)		-0.206*** (0.0535)
NONPROFIT	-0.0611*** (0.0276)		-0.128*** (0.0432)
Payor Distribution			
%PRIVATE		-0.000772** (0.000351)	-0.000268 (0.000529)
%MEDICARE		0.00255 (0.00238)	0.00696* (0.00354)
Survey Certification Ratings			
NI		0.0294 (0.0185)	0.0302 (0.0289)
DEFERRED	0.0550*** (0.0203)	0.0258 (0.0182)	0.101*** (0.0284)
#UR		0.00136 (0.00118)	
Bed-Size Categories			
SIZE2		0.0133 (0.0114)	
SIZE3			0.0193 (0.0169)
SIZE4	0.0246 (0.0189)		
SIZE5			

*Significant at .10 level.
**Significant at .05 level.
***Significant at .01 level.

Table T-5.6 (cont'd.)

Independent Variable	1974-1975	1975-1976	1976-1977
Regions			
REGION2		0.0144 (0.0181)	0.0197 (0.0312)
REGION3	-0.0162 (0.0166)	-0.0195 (0.0153)	-0.0500** (0.0249)
REGION4			-0.00707 (0.0311)
REGION5	-0.0381* (0.0192)		-0.0592** (0.0248)
REGION6			
REGION7	0.0252 (0.0241)		
TECHNOLOGY OR PRODUCTION PROCESS VARIABLES			
SNF-HRF	0.0326*** (0.0114)	0.0257** (0.0124)	0.0621*** (0.0193)
HRF-SNF	0.158* (0.0875)		0.338*** (0.119)
CLINIC			0.0551 (0.0547)
%CONTRACT			
FACILITY AGE	-0.000310 (0.000310)		-0.000595 (0.00043)
INDICATORS OF SHORT-RUN DISEQUILIBRIUM			
Occupancy Segments			
OCC1		0.000860 (0.00081)	0.00137* (0.00073)
OCC2			
OCC3			
OCC4			
CONSTANT	0.342	0.202	0.228
Summary Statistics			
Coefficient of Determination (R^2)	0.362	0.405	0.432
Corrected Coefficient of Determination	0.232	0.297	0.283
F-Ratio	2.77	3.76	2.89

*Significant at .10 level.
**Significant at .05 level.
***Significant at .01 level.

is not significant in either the 1975-1976 or 1974-1976 models. (It was inadvertently omitted in the 1974-1975 model.) The ΔMEDICARE variable obtains positive coefficients in all three years, but is significant only for 1974-1976. The ΔADMRATE variable gives positive and significant coefficients for both 1974-1975 and 1974-1976. The ΔFEMALE coefficients are positive in all three models, but only the one for 1975-1976 is significant. Although the ΔWALK coefficient is positive in the 1974-1975 model, it is negative in the 1975-1976 and 1974-1976 models. In conclusion, these other first-difference results have been generally disappointing. The ΔRATE results are much more robust by comparison.

The other "absolute level" variables in these first-difference models were included largely to control for other relevant differences and the results are examined only briefly. First, there is considerable inconsistency between the various models in the sense that different variables enter the different models and do so with differing significance and occasionally different signs. Nevertheless, a rather interesting pattern of results emerges for these variables. In particular, for many variables included in each of the three models the coefficients are reversed from those predicted and obtained in cost models for the various years (see Section T-5.3). For example, the 1975-1976 coefficient for %WALK is negative and significant; the coefficient for DRUGS is negative and significant; the coefficient for #X-RAYS is negative and significant; the coefficient for %3BED is positive and nearly significant; and the coefficient for LPN WAGE is negative and significant. That is, facilities with the most costly patients and service-mixes and highest labor costs are experiencing lower rates of cost increase. Of course, this is not what would have been predicted. A possible explanation is the following. These factors may reflect facilities having higher costs, ceteris paribus, and may therefore also proxy facilities "bumping up" against the ceiling limitations for reimbursement. If so, the results signal that the ceiling limitations have been effective in moderating the rate of cost increase for such higher cost facilities. This result raises another concern. If this explanation is correct, it signals that rate setters in New York are discriminating in favor of those facilities providing fewer services and serving a less costly patient-mix. Although this result should not be surprising, given that such differences are not explicitly

taken into account in rate setting, it is not apparent a priori that such a result is desirable.[1] It provides an incentive to reduce the service intensity and, presumably, thereby lower the quality of care. If this were indeed an objective of rate setting, it could be more effectively promoted in some other fashion--e.g., non-reimbursement of costs associated with certain services. However, it is more plausible that such behavior is an unintended side effect of a comparatively simplistic approach to rate setting. As in other states, the reimbursement authorities in New York have been unwilling to make explicit decisions about the "quality" differences for which they are willing to pay or about the differential resource requirements associated with patients having differing needs. It has perhaps been assumed that such relevant differences are proxied by certification and other grouping criteria. If so, the results suggest that such control has not been adequate.

Finally, is noteworthy that PROFIT and NONPROFIT were significant in the 1974-1975 model but did not even enter with coefficients so large as their standard errors in the 1975-1976 model. This may reflect the effect of discontinuing to group by sponsorship in 1976. (Of course, it may also reflect coefficient instability; a situation which can be observed with respect to other independent variables.)

Although the first-difference approach in contrast to a pooled cross-section/time-series analysis has been helpful in some respects, it has been disappointing overall. It seems to have done little more than to indicate or hint at the cost and other consequences of a rather unsophisticated procedure for rate setting. Although the findings from this analysis may argue the case for a formulary system, they do not in any way contribute to the design of such a system. Cross-sectional analyses such as those reported in Section T-5.3 offer a much more meaningful basis for this purpose. After all, one should have a primary focus on setting appropriate rate levels and not on the changes occuring from year-to-year. An inflation factor can be used for the latter purpose.

T-5.5 Summary and Discussion of Policy Implications

Summarizing all of the findings reported above would be a much too arduous task, both for the reader and the author. Rather, this section emphasizes those findings that have larger significance and implications for the design of a formulary prospective reimbursement system.

---

[1] Such differences are no doubt partly reflected in the grouping criteria.

First, the analysis indicates that "scale" and occupancy are minimally-important to determining cost variation. Although increasing returns are indicated between 0 and 50 beds and diminishing returns above 300 beds, these relationships are neither particularly significant nor dramatic. Although analyses of 1974 and 1976 data found cost levels to decrease significantly between 80 and 90 percent occupancy, the more dependable analysis using 1975 data found no meaningful relationship between cost and occupancy. It appears that regulators have been too concerned with minimally-important scale factors at the expense of other relevant differences that have much more dependable and quantitatively-important impacts on the cost level.

Second, patient- and service-mix differences do make a difference. The New York State Moreland Commission had concluded from analysis of the substantially limited data available from 1973 cost reports that "patient characteristics and patient needs by and large do not significantly explain variations."[1] Our analyses of the more recent and more exhaustive 1975 and 1976 cost report data do not support this conclusion. Although patient age and sex variables are neither statistically significant nor quantitatively important, the patient functional status variables (e.g., percentage of patients requiring assistance with eating) are highly significant as a group and very important in terms of explaining the pattern of nursing home cost differences. Service intensity variables (e.g., number of physical therapy sessions were similarly significant and important. In addition, several service offering and patient impairment (e.g., percentage of patients with sight impaired) variables were also statistically significant. Moreover, our analysis of year-to-year cost changes indicates that failure to take account of these cost factors in rate setting implies unfair or inequitable budgetary stringency with regard to facilities that provide more services and have a more expensive patient mix.

Third, the sponsorship, certification and region variables were not only highly significant but also gave comparatively large coefficients. Having estimated some of the most ambitious and general cost functions to

---

[1] Reimbursing Operating Costs: Dollars Without Sense, p. 54.

date, it is nevertheless found that nonprofit facilities in New York cost four to six dollars more per day than otherwise similar for-profit facilities, that SNF's cost $9.00 to $11.00 more per day than HRF's after controlling for service- and patient-mix differences, and that facilities in the New York City area cost $4.00 to $5.00 more than other facilities in the State of New York after controlling for regional and even facility-specific wage differentials. The question needs to be asked: What accounts for these differences if not differences in the patient mix, service mix, or input prices? Surely, differences of this magnitude cannot continue to be reimbursed without knowing what causes them. If they are merely due to unmeasured differences in the amenity level of care or to managerial inefficiency, then it is not clear that public programs should continue to discriminate between facilities in these categories and pay the differences.

APPENDIX T-5.A

NEW YORK STATE NURSING HOME REIMBURSEMENT SYSTEM

APPENDIX T-5.A

NEW YORK STATE NURSING HOME REIMBURSEMENT SYSTEM*

Between 1971 and November 1976, New York's Prospective Reimbursement (PR) Regulations on Nursing Homes were not changed significantly. Several important changes, however, were made in 1976 and further changes that were made in 1977 took effect in January 1978. The specifics of many of the regulatory changes are set forth in the State's annual update of Part 86 Reporting and Rate Certification for Medical Facilities. The following discussion will highlight the important nursing home reimbursement regulations.

Until November 1976, nursing homes (SNFs) and health related facilities (ICFs) were grouped by three characteristics: geography (seven groups), bedsize (five groups) and sponsor (three groups). If a group did not have a sufficient number of facilities to establish a reimbursable ceiling, such facilities were combined into another similar group or combined with comparable medical facilities without regard to geographic areas or size. In 1976 sponsorship was eliminated as a grouping category; the geographic grouping was changed to reflect geographic wage differentials. Eliminating sponsorship meant that public and voluntary facilities may be grouped with proprietary ones. This provision has upset the higher cost public and voluntary facilities. While the percentage of facilities hit by reimbursement ceilings caused by this change in grouping categories is not known, it is perceived to be higher than the 15 percent which is alleged to have hit the ceiling prior to this regulatory change.

Due to the outcry from nursing homes and health related facilities caused by the grouping changes, New York State instituted a mechanism to waive the reimbursement ceiling. The State sent in a team to review the intensity and scope of services of facilities requesting a waiver. If the team felt that the case mix justified the services, a waiver was granted. The review process was not implemented until 1977. The 1976 data therefore should not reflect changes in 1976 revenue as a result of any decision in 1977 to waive the reimbursement ceiling.

---

*by Michael Sumner

The reliability of the 1976 revenue data is also adversely affected because the regulatory changes delayed rate setting until November 1976. Nursing homes until then were reimbursed at their 1975 rate. Because the 1976 regulations were made retroactive to January of that year, subsequent revenue adjustments may cause 1976 revenues to be restated.

The confusion surrounding the 1976 rate has resulted in litigation such that no nursing home has annual revenues in 1976 based on their 1976 rate. Many nursing homes (probably the non-profit ones) received lower rates in 1976 than in 1975. They have not returned to the state the difference owed nor has the state paid the differential to nursing homes with higher rates in 1976 than in 1975.

The methodology to compute reimbursement rates for each nursing home is complex. In computing the reimbursement rate for nursing homes, the cost of administrative, dietary and housekeeping services for all homes in a group cannot be in excess of a ceiling applied to not more than 10 percent of the total nursing homes in the state. In no event do the total allowable costs, excluding property costs, costs for therapy and prescription drugs, return on net equity capital and incentive allowance, exceed 115 percent of the weighted average cost of all nursing homes in the group.

For nursing homes which are affiliated with and operated in conjunction with a hospital, the maximum allowable costs which will be included in computing the reimbursement rate may be no more than 40 percent of the allowable per diem operational costs of the related hospital. This provision may be waived by the Commissioner of Health when the facility can substantiate higher nursing home costs due to a short length of stay and/or level and intensity of care provided. For ICFs affiliated with SNFs the maximum allowable costs which will be included in computing the reimbursement rate may be no more than 60 percent of the allowable per diem operational costs of the affiliated SNF. For ICFs which are not affiliated with another certified facility, the maximum rate may not exceed 60 percent of the weighted average per diem operational costs of nursing homes of similar size in the same geographic area. For these ceilings, operating costs are defined as total costs less depreciation on real property, real estate taxes, interest on real property capital indebtedness, insurance on buildings, and real property rental payments.

To the allowable basic rate computed in accordance with ceiling limitations, there is added an inflation factor. Such factor is determined by weighing nursing homes cost elements. The elements are:

Salaries and employee health and welfare expense to be computed on a total of sub-weights

Non-payroll administrative and general expense

Non-payroll dietary expense

Non-payroll professional care expense

The weight for each element is adjusted by an appropriate price index.

Prior to 1975 any proprietary nursing home was eligible to receive an incentive allowance for a portion of the amount, as established by the Commissioner of Health, by which its total cost of administrative, dietary and housekeeping services were lower than the weighted average of the costs of these services for the group. This incentive was withdrawn in 1975.

Until November 1975 labor costs were subject only to internal industry wage movements. The method only penalized those facilities with hourly labor costs significantly higher than its group's average. Labor costs, in effect, were not subject to a ceiling. Now, the reimbursement rate includes a wage index tied to external wage movements. With labor costs approximately 70 percent of total costs, facilities will have to contain labor cost more so that previously.

The above mentioned reimbursement method applies to operating costs. Capital costs are not subject to the ceiling on operating costs. Prior to January 1975 facilities were reimbursed for their depreciation.[1] Depreciation methods were allowed to differ for facilities of different sponsorship. Now the depreciation method has been made consistent for voluntary and public facilities. Depreciation on plant and non-movable equipment must be based on the straight line method. Depreciation on movable equipment can be either straight line or accelerated.

---

[1]Older proprietary facilities with very low capital costs (i.e., depreciation and interest) could have their reimbursement rate based on an imputed rent rather than depreciation and interest. The intent of this procedure was to allow older facilities to survive and make the capital improvements that would be necessary to retain their operating licenses.

Proprietary facilities are not reimbursed for their capital costs. Instead reimbursement is based on a fair rental value. There have been many problems with the use of fair rental values. Both the state and the facilities admit that the concept has not worked. Apparently some facilities have gained revenue but more have received less than their incurred interest and depreciation cost, and thus do not meet their perception of a fair rental value.

Since the inception of New York's PR system, a facility's reimbursement rate could be subject to penalities other than the group ceiling on operating costs. If a facility was late in submitting required financial and statistical reports, the Commissioner of Health could reduce the rates until the reports were submitted. Prior to 1976, a facility could have its rate reduced ten percent. Now rates are reduced only two percent. The penalty for facilities with substantial deficiencies has also changed. The original penalty would not reimburse a facility above its group weighted average if the facility had substantial operational deficiencies. In November 1976 this penalty provision was changed. Now a facility's reimbursement is affected in rough proportion to its ability to satisfy the survey/certification standards. If a facility exceeds state and federal standards, the facility is not subject to the reimbursement penalty. If a facility complies with but does not exceed state standards, its reimbursement is limited to, at maximus, the 60th percentile of costs for comparable facilities which comply with the state standards. If a facility does not meet federal standards in one of four major areas, its reimbursement can be no greater than the 10th percentile of costs for comparable facilities which comply with the state standards. The penalties, however, could be waived if a facility is able to tie its higher costs to a more complex case mix.

The 1976 changes have not worked. This year the penalties have been again changed. Facilities are now ranked into three quality categories: very good, satisfactory, and needs improvement. If a facility is judged to be very good, its reimbursement ceiling will be the higher of its group's weighted reimbursement rate or 110 percent of the weighted average of the comparable satisfactory group. If a facility is judged to be

satisfactory, its ceiling will be the group's weighted average rate. If a facility needs improvement, its reimbursement ceiling will be the lower of its group's weighted average or 90 percent of the comparable satisfactory group.

There has been one other important reimbursement change. There is no mid-term rate adjustment. If actual costs exceed previously projected costs, any rate change granted by appeal or blanket addition similar to the fuel increase in 1974 will be added to the reimbursement rate in the following year. The regulation which became effective in 1976 attempts to force nursing homes to live within their projected costs.

The remaining regulations are fairly typical of other states. Profit is calculated using the Medicare methodology for return on equity. Charity and courtesy allowances, purchase discounts and other allowances are not reimbursed. Tuition and supporting grants are allowable costs as are patient care services and bad debts resulting from failure of Medicaid patients to pay the co-payment for outpatient services.

Finally, it must be noted that New York has not adopted a uniform chart of accounts and definitions. Nursing homes categorize costs as they wish to.

APPENDIX T-5.B

MEANS AND STANDARD DEVIATIONS OF NEW YORK DATA

Table T-5.B.1: Means and Standard Deviations for 1974 (479 cases)

Variable	Mean	Standard Deviation
Scale		
BEDS	117.088	96.794
Utilization		
OCC	0.944	0.115
Patient Age Distribution		
Percent Under 65	5.8574	6.0510
Percent Age 75 to 84	41.0442	9.2482
Percent Over 84	38.0083	12.6432
Sex		
Percent Female	76.0450	12.0064
Patient Functional Status		
Percent Walk	61.6808	26.2528
Case Flow		
Percent Deaths	25.0119	17.5250
ADM Rate	0.0028	0.0018
Area Descriptors		
RETAIL WAGE	2.7722	0.3112
POPDEN	10741.5595	17934.0928
Location		
Region 2	0.0898	0.2862
Region 3	0.1336	0.3406
Region 4	0.1106	0.3140
Region 5	0.1441	0.3515
Region 6	0.1211	0.3266
Region 7	0.2839	0.4514
Facility Characteristics		
NON PROFIT	0.2860	0.4524
PROFIT	0.6117	0.4879
PCT CARE	1.7086	3.3427
PCT PRIV	24.9440	22.5865
NI	0.0755	0.2369
DR	0.0686	0.2292

Table T-5.B.2: Means and Standard Deviations for 1975 (504 cases)

Variable	Mean	Standard Deviation
Scale		
BEDS	123.323	97.743
Utilization		
OCC	0.966	0.075
Patient Age Distribution		
Percent Under 65	6.2150	6.4831
Percent Age 75 to 84	41.0777	8.8731
Percent Over 84	37.9258	13.1485
Sex		
Percent Female	75.4988	11.5115
Patient Functional Status		
Percent Walk	64.7220	25.9201
Percent Eat	45.5447	29.1556
Percent Bathe	89.4055	18.4965
Percent Breathe	24.9556	41.5806
Percent Dress	75.1456	29.5279
Percent Correspond	68.9165	32.1450
Percent Medicate	94.8464	15.9897
Percent Massage	81.9314	32.5681
Patient Impairment		
Percent Alert		
Percent See		
Percent Hear		
Percent Speak		
Percent Communicate		
Case Flow		
Percent Deaths		
ADM Rate	0.0029	0.0032
Service Offerings (0-1)		
Drugs	0.6607	0.4739
Phys Therapy	0.8651	0.3420
OCC Therapy	0.6290	0.4836
Speech Therapy	0.4365	0.4964
Special Duty	0.0337	0.1807
Phys Services	0.5139	0.5003
Other Services	1.6111	1.2782

Table T-5.B.2: Means and Standard Deviations for 1976 (504 cases)
(continued)

Variable	Mean	Standard Deviation
Service Intensity		
No. Prescriptions	34.5136	88.6568
No. X-Rays	0.7824	2.5387
No. Physical Therapy	0.2954	0.3522
No. Speech Therapy	0.0178	0.0486
No. OCC Therapy	0.0840	0.1906
No. INHAL Therapy	0.0072	0.0373
No. Radiation Therapy	0.0048	0.0638
No. SOC Service	0.6586	1.0342
Bedroom Distribution		
Percent 1 Bed	21.6757	23.1621
Percent 3 Bed	18.7613	26.5496
LPN WAGE	5.2519	1.5536
RETAIL WAGE	2.7562	.2915
POPDEN	10325.7063	16456.0915
Regions		
Region 2	0.0933	0.2911
Region 3	0.1349	0.3420
Region 4	0.1190	0.3242
Region 5	0.1349	0.3420
Region 6	0.1032	0.3045
Region 7	0.3056	0.4611
Facility Characteristics		
SNF-HRF	0.2361	0.4251
HRF-SNF	0.2143	0.4107
CLINIC	0.0159	0.1251
Percent CONTRACT	3.5947	5.0032
FACILITY AGE	34.7408	35.1415
HRF	0.2837	0.4513
PROFIT	0.6091	0.4884
NON PROFIT	0.2837	0.4513
NI	0.0788	0.2496
DF	0.0639	0.2282
PCT PRIV	22.3778	20.5891
PCT CARE	1.4103	2.8601

Table T-5.B.3: Means and Standard Deviations for 1976 (497 cases)

Variable	Mean	Standard Deviation
Scale		
BEDS	125.436	96.213
Utilization		
OCC	0.966	0.081
Patient Age Distribution		
Percent Under 65	5.6942	5.8698
Percent Age 75 to 84	40.6037	8.5864
Percent Over 84	40.0218	12.5212
Sex		
Percent Female	77.1544	10.2128
Patient Functional Status		
Percent Walk	61.2283	27.6872
Percent Eat	46.9415	29.8093
Percent Bathe	91.4810	16.6294
Percent Breathe	23.9582	41.1427
Percent Dress	75.1456	28.9788
Percent Correspond	71.3462	30.2546
Percent Massage	87.3951	28.4657
Patient Impairment		
Percent Alert	59.3705	27.2837
Percent See	63.2829	29.8427
Percent Hear	45.7704	29.9065
Percent Speak	28.2209	25.7996
Percent Communicate	38.5696	27.9299
Case Flow		
Percent Deaths	19.4276	15.5445
ADM Rate	0.0024	0.0012
Service Intensity		
No. Prescriptions	24.2126	36.0435
No. X-Rays	0.8905	2.9579
No. Physical Therapy	0.2690	0.4625
No. Speech Therapy	0.0153	0.0352
No. OCC Therapy	0.0907	0.7550
No. INHAL Therapy	0.0048	0.0210
No. Radiation Therapy	0.0012	0.0085
No. SOC Service	0.5811	0.6168

Table T-5.B.3: Means and Standard Deviations for 1976 (497 cases)
(continued)

Variable	Mean	Standard Deviation
Bedroom Distribution		
Percent 1 Bed	21.4047	16.9864
Percent 3 Bed	66.1855	26.5768
RETAIL WAGE	2.7557	0.2845
POPDEN	10052.6720	15613.4011
Regions		
Region 2	0.1207	0.3261
Region 3	0.0946	0.2929
Region 4	0.1247	0.3308
Region 5, Region 6]*	0.2555	0.4366
Region 7	0.3058	0.4612
Facility Characteristics		
SNF-HRF	0.2515	0.4343
HRF-SNF	0.2455	0.4308
HRF	0.3078	0.4621
PROFIT	0.6076	0.4888
NON PROFIT	0.3119	0.4637
NUMUR	3.8109	3.3046
PCT PRIV	21.9787	20.1504
PCT CARE	2.2405	3.1674
NI	0.0724	0.2595
DF	0.0584	0.2346

*Regions were combined for grouping purposes.

TECHNICAL PAPER #6

Technical Paper #6

NURSING HOME COSTS IN MASSACHUSETTS*

T-6.0 Introduction

This technical paper presents each of the three econometric cost functions estimated for facilities in the Commonwealth of Massachusetts. Like the other two state-specific sections of this project (for New York and Indiana), it can be used to indicate the type of results Medicaid (i.e., state-oriented) rate setters would find in applying the formula prospective reimbursement methodology presented by this project. These state-specific results can also shed light on the national cost function results, since there is no variation in regulations within a state cross section. The sensitivity of national cross section results to variation in regulation can be evaluated by examining the consistency of national and state findings.

The remainder of this chapter is organized as follows. Section T-6.1 reviews the Commonwealth's nursing home reimbursement system for 1973, 1974 and 1975. These were the three years for which data were analyzed. Section T-6.2 briefly reviews the conceptual model developed more fully in Technical Paper #3. Section T-6.3 operationalizes the model in accordance with characteristics of the Massachusetts nursing home industry and the available data. In this section, specific hypotheses regarding the nature of per diem costs are developed. Section T-6.4 discusses data sources, data definitions, data consistency, and the treatment of missing observations. In Section T-6.5 the econometric cross-sectional results are discussed in detail. Sections T-6.6 and T-6.7 present the results of additional analyses, a first difference analysis of operating expenses and an analysis of selected components of operating expenses. The final section, Section T-6.8, summarizes the results and explores the policy implications of the findings.

T-6.1 The Nursing Home Reimbursement System in Massachusetts

Only minor changes occurred in the nursing home reimbursement system of Massachusetts between 1973 and 1975. The section below presents an overview of the system operating throughout this period.[1]

---

*by Christine Bishop.

[1]Appendix T-6.A gives a more detailed report of the system.

The nursing home reimbursement system operated by the Massachusetts Rate Setting Commission (RSC) from 1973 through 1975 had both prospective and retrospective elements. In the process of setting a rate structure, nursing homes were grouped by two characteristics: certification (four categories) and bedsize (two categories). An individual provider's interim reimbursement rate was set equal to the per diem operating expenses for that nursing home adjusted by an inflation factor with two important exceptions. First, limitation on nursing home costs provisions established a ceiling for each category of nursing home. The maximum allowable interim reimbursement rate for any nursing home was therefore equal to the average cost for the respective group plus one standard deviation. Second, an occupancy penalty was incorporated into the system as follows. The interim rate was the sum of several operating cost components computed on a per diem basis. To incorporate an occupancy penalty in the setting of the rates, interim rates for some components were established by dividing total allowable costs or expenses by the larger of (1) the facility's actual patient-day census for the year of historical data or (2) a fixed percentage (90% used in 1973 and 1974, and 93% used in 1975) of the facility's total bed capacity.

The final rate of payment was determined in much the same way as the interim rate except that actual expenses determined by a field audit were used instead of historical costs. Under certain circumstances (discussed in Appendix T-6.A), an operator could have the final rates adjusted via an appeals process.

T-6.2 Reviewing The Theoretical Framework For Cost Function Analysis

The average cost functions to be estimated for Massachusetts are similar in general form to those estimated for a national cross section of facilities as discussed in Technical Paper #4. To review, average operating cost per patient day is hypothesized to be determined by a function which can be stated in general terms as the following.

$AC = f(Q, S, X, q, A, P, M, R, Y)$

where

Q = quantity of output produced

S = services provided per patient day

X = patient mix

q = quality of care

P = input prices

R = regulation

Y = additional facility characteristics

The "output" or "product" of nursing homes cannot be meaningfully measured as a single quantity. Whereas patient days of care is a useful indicator of the output level, it it only one of many dimensions describing relevant differences between the outputs of different facilities. In particular, the number of patient days per se reflects neither variation in the "quality" of care nor variation in the characteristics of patients receiving care--e.g., it is likely to be more expensive, other things equal, to care for a nonambulatory patient than for an ambulatory patient. Because nursing homes do not produce a standardized product it is important that differentiating aspects of the product be captured in the cost function. In Technical Paper #3 it is argued that service intensity, in other words the services available and offered to patients, is what differentiates the nursing home product. Quality of care is then determined by the appropriateness of this care to patient mix. Patient functional status and diagnosis may also have an impact on cost, independent of what is available to patients in the facility. Quality of care, for which no adequate indicators are available, can also be seen as a determinant of cost: providing high quality care to a given mix of patients is generally assumed to be more costly than providing low quality care to the same mix. Finally, case flow may be seen as a factor that differentiates the product of the facility, although it is also an indicator of scale of production when product is measured in cases treated, and as an indicator of the intensity of use of the facility capacity.

The specific variables used to represent each of these general groups of variables is discussed in turn below.

T-6.3 Operationalizing the Conceptual Model

T-6.3.1 Choice of Variables

Theoretical considerations indicate variables that should be included in an average cost function, and the findings of others (see Technical Paper #4) also recommend inclusion of certain types of variables. Unfortunately, not all of these variables are available for Massachusetts for all three periods considered here, calendar years 1973, 1974, and 1975. Thus the ideal variables must be proxied by available variables.

Scale. Because of Massachusetts regulation of capacity under certificate of need, capacity of facilities can be assumed fixed for purposes of their decision making. In a cross section, then, profit maximizing firms are not necessarily observed to be operating on the long run cost function (i.e., producing their chosen output at long run minimum cost for that output). However, they can be assumed to be operating on the short run average cost function for their bed size. This leads to the inclusion of both the scale of capacity and utilization of capacity in the cost function. The variable, beds, shows the shift from one short run average cost curve to another as capital investment is allowed to vary; the variable, average occupancy rate, allows utilization to vary given capacity. A total output variable would be total patient days provided. This is equal to occupancy rate times bed times a constant:

$$\text{patient days} = (\text{occupancy}) \times (\text{beds}) \times 365$$

In the Massachusetts data, scale of production is represented by the mean operating beds in a facility for the period and by the average occupancy rate (total patient days divided by available bed days). An alternative to the beds variable is average daily census (patient days divided by 365).

Service Intensity. Service intensity can be represented in all three years by the proportion of beds licensed to provide care at a Skilled Nursing Facility (SNF) level or at an Intermediate Care Facility (ICF) level, and also by the proportion of rest home beds. (Free standing rest homes were

not included in the data set.) In Massachusetts, beds certified to provide Medicare services must meet higher input standards than beds licensed to provide SNF care. For this reason, it is desirable to include the proportion of beds certified for Medicare and the proportion that are SNF but not certified for Medicare. This is possible for 1974 and 1975; in 1973, only a dummy for Medicare certification is available. In general, these certification and licensing proportions ensure that a common minimum service intensity is available to patients occupying those beds, no matter how this care is actually paid for. Facilities may provide more than these minimum standards if they so choose.

Of course it would be preferable to include service availability variables directly rather than, in essence, relying on state inspectors to define what the nursing homes are providing. This is possible for the national data, and means that in the national analysis certification and licensure status can be used as regulatory variables. Since in Massachusetts all facilities must be licensed before care can be provided to anyone, and since there is no evidence that the standards for SNF beds and ICF beds are differentially enforced, use of these bed proportion variables to indicate differential regulation is not appropriate. They are, for better or worse, the best indicators present in the data of service availability.

The availability of certain services, whether utilized or not, can be indicated for all three years. For 1973, dummy variables from a questionnaire show whether physical therapy, occupational therapy, speech and hearing therapy, and social services were available to residents. (Other services are also indicated, but the proportion of facilities providing them is so high as to make them virtually standard.) For 1974, whether physical therapy (PT) was available was determined by whether facilities reported any costs for PT; in 1975, a non-zero value for the number of PT hours employed was taken as an indication of PT availability and the dummy variable, PT, was set to 1 for these facilities. In either year, a physical therapy dummy variable was set equal to 1 if availability could be inferred.

A variety of variables are suggested by the health cost function literature for capturing the case flow through a facility. Admissions per occupied bed is used by Feldstein (1968); an alternative is admissions per patient day (ADMISPD), which differs by a factor of 1/365 and has the

property that the addition to total cost of one more admission (holding all else constant, including patient days) is shown by the coefficient of the variable in the linear cost function.[1]

Input Prices. Input prices shift the cost function and must be included explicitly or proxied by location variables.

Data are available from the Area Resource File for the retail wage bill per retail employee for the 12 counties in Massachusetts. Regional dummies for the eight health planning regions (or the six Health Systems Areas in 1975) can be used to proxy price variation. Population density, available by county, can be used to indicate urban-rural input price differentials.

Regulation. Regulations do not vary within one state, but the impact of regulations may well vary across facilities. Intuitively a facility serving many Medicaid or other public patients will be more affected by Medicaid rate setting regulations than are facilities which serve more private pay patients. Facilities with a high proportion of Medicaid patients may be more likely to provide a service intensity that has costs below state reasonable cost ceilings, and achieve the state occupancy floor (90%) in order to avoid a rate penalty. Proportion of patients by payment type on the census date will be used to indicate the degree to which facilities are subject to rate setting restriction.

Regional dummies are used to account for differential regulation in the national analysis. In Massachusetts, despite some evidence of regional variation in input prices, location does not play a part in rate setting, and regulation is carried out on a centralized, not a regional, basis. Thus regional dummies are not expected to pick up regulatory variation in the Massachusetts analysis.

---

[1] Length of Stay (LOS) is a variable that can be used as an alternative to admissions rate. It is approximately equivalent to the inverse of admissions divided by patient days, assuming that admissions are approximately equal to discharges and have occurred at a relatively stable rate over a long period in any facility. Thus the coefficient of length of stay is expected to have a negative sign if admission are costly ceteris paribus. Variables describing the distribution of length of stay may better describe what happens to patients in facilities. The proportions of discharged patients within different length of stay categories are available for Massachusetts.

Additional Facility Characteristics. Certain characteristics of a facility may affect the operating cost. Ownership is a determinant of product attributes but may also be seen as a variable describing ability to minimize costs, if non-profit operators are less aggressive about achieving efficient production (i.e., if they tend to operate above, not on, their cost curves).[1] A dummy variable, NONPROF, was included in the model, and distinguished facilities which are nonprofit. A dummy variable, CHAIN, was included indicating whether or not a facility was part of a chain. The chain variable is not available for 1973; for 1974, joint operation is indicated by a dummy variable; and for 1975 information is available for the number of other facilities operating under the same ownership. This last, the continuous variable NCHAIN, can be included quadratically to allow both economies and diseconomies of ownership scale.

Another facility characteristic is represented by the variable MIXED, which equals 1 if the facility is licensed to provide two levels of care under the same roof. Unlike the data for New York, the observations in the Massachusetts data set are for complete facilities, which may be freestanding SNFs or ICFs, or may combine several levels of care (including rest home care). It is possible that the production function, and hence the cost function, for production of mixed levels of care is not well represented by a weighted average of the effects of the impacts of each of the three levels, which is included as proportion of beds at each level. A facility providing any SNF care may provide a higher intensity of care to its ICF patients through a spill-over effect.[2]

---

[1] The counter argument is that optimizing a non-profit objective function still requires cost minimization. See Scanlon (1978) for a treatment of non-profit optimization demonstrating this point.

[2] There is also a regulatory effect at work for these facilities in Massachusetts. Only one rate is set for mixed facilities, determined using average cost screened by a weighted average of the ceilings for freestanding SNFs and ICFs. The apparent incentive is to produce relatively expensive ICF care for public patients at the high rate mixed (relative to freestanding ICF beds) and to attract private patients to fill the remainder of the high intensity ICF beds and all SNF beds. It is not known whether this occurs, but it is not surprising that average costs for mixed facilities are higher than a comparable weighted average of freestanding SNF and ICF facilities.

The service provided by plant and equipment may be complementary with operating services provided. The services of capital are represented by the probability that a patient day is provided in a single room, which is measured by the proportion of beds in one bed room for 1974 and 1975. The proportion of beds in three and four or more bed rooms is also used.[1]

T-6.3.2 Hypotheses To Be Tested

A number of hypotheses can be developed with respect to these variables. These are discussed in general in Technical Paper #3 and in Bishop, Reagan and Birnbaum (1977).

Scale

1. There are economies of scale for nursing homes, so that firms providing a larger output (patient days) have lower average costs.
2. There are economies of utilization: given capacity, average operating cost falls as occupancy increases. This implies that a significant proportion of operating cost is relatively fixed.

Output Characteristics

1. Higher service intensity is associated with higher nursing home costs.
2. Patient mix does not affect average cost once service intensity is accounted for.
3. The rate of admission per patient day is positively associated with average cost.

Input Prices

Facilities located where input prices are high have higher average costs.

Regulation

1. Service intensity, and hence average cost, is higher the higher is the proportion of Medicare patients served.
2. Costs are lower the higher is the proportion of Medicaid patients served due to implicit ceilings on service intensity.

Facility Characteristics

1. Non-profit facilities do not have higher average costs once service intensity is accounted for.
2. Facilities operated as part of a chain have lower average costs.

---

[1] It should be noted that the coefficient on these proportion variables in the additive equation can be interpreted as the incremental cost of providing care to a patient in a single, double, or triple room, or a ward. Other researchers (AMS, 1976) have used proportion of rooms by type. Although this also describes capital endowment, it is not so easy to interpret, because it is not possible to change the proportion of rooms of one type by one unit while holding bed size constant.

T-6.4 Data

Data describing expenses and facility characteristics for Massachusetts nursing homes for 1973, 1974, and 1975 came from two sources. Expense data must be reported on form RSC-1 for all facilities requesting a rate for care of public patients. These are accounting costs, and so are subject to allocation conventions of individual respondents; the subcategories of operating expense are not well delineated for this reason. They are also affected by the purpose for which the data are collected: it is possible that some facilities may under-report costs which they know are not reimbursable in the rate, and since these data come from unaudited cost reports, they may also contain overreporting in some categories.

Patient data by payment source also appears on the RSC-1. The second source of data is a survey of facilities conducted by the Massachusetts Department of Public Health each spring. It contains information on facility characteristics like ownership, capacity by level of care, and location; operating statistics (patient days by level, admissions and discharges by type, some employment data) from the previous year; and information from a one-day census of patients, classifying them by diagnosis, age, sex, and payment source. Files which merged these two data sets for each of three years were supplied by Boston University School of Management Health Care Management Program.

Several types of facilities had incomplete or erroneous records and/or were believed to be non-representative for purposes of examining nursing home cost behavior. Facilities that opened, closed, or changed ownership during the year reported only past-year cost data. Facilities that expanded or contracted their bed size by more than 10% were believed to be of interest for a study of facility expansion, but non-representative for determination of ongoing average costs. Finally major discrepancies between patient days as reported on the RSC-1 and patient days reported on the survey called this key variable into question. If patient days (PD) from these data sources differed by more than 10%, the observation was eliminated. Freestanding rest homes were omitted as beyond the scope of this study.

Some facilities were also dropped because of the absence of other important variables (census, admissions, discharge). The variables used for analysis and their respective code names are reported in Table T-6.1.

Table T-6.1

VARIABLE DEFINITIONS, MASSACHUSETTS ANALYSIS

AOC	Average Operating Cost Per Patient Day
Scale	
Capacity	
BEDS	Licensed bed capacity, average for year; see footnote 1
Utilization	
OCC	Total pationt days divided by total available bed days (=BEDS x 365) times 100; see footnote 2

1
To operationalize the grafting procedure, bed join segments were defined as follows:

BED0	= BEDS if BEDS $\leq$ 40
	= 40 if BEDS > 40
BED40	= 0 if BEDS $\leq$ 40
	= BEDS - 40 if 40 < BEDS $\leq$ 80
	= 40 if BEDS > 80
BED80	= 0 if BEDS $\leq$ 80
	= BEDS - 80 if BEDS > 80

2
To operationalize the grafting procedure, occupancy join segments were defined as follows:

OCC0	= OCC, if OCC $\leq$ 90
	= 90, if OCC > 90
OCC90	= 0, if OCC $\leq$ 90
	= OCC - 90, if 90 < OCC $\leq$ 93
	= 3, if 93 < OCC
OCC93	= 0, if OCC $\leq$ 93
	= OCC - 93, if 93 < OCC $\leq$ 95
	= 2, if OCC $\leq$ 95
OCC95	= 0, if OCC $\leq$ 95
	= OCC - 95, if 95 < OCC $\leq$ 97
	= 2, if OCC > 97
OCC97	= 0, if OCC $\leq$ 97
	= OCC - 97, if OCC > 97
OCGE90	= 1, if OCC $\geq$ 90
	= 0 otherwise
OCGE93	= 1, if OCC $\geq$ 93
	= 0 otherwise
OCGE95	= 1, if OCC $\geq$ 95
	= 0 otherwise

Table T-6.1 (Continued)

Patient Proportion	
Age	
PLE49	Proportion of patients on census date who were less than or equal to 49 years old
P5064	Proportion between 50 and 67 years old
PLE59	Proportion 59 or less
P6569	Proportion between 65 and 69
P8089	Proportion between 80 and 89
POLD	Proportion 90 years old or more
Sex	
PFEM	Proportion of patients on census date who were female
Admissions Source	
PADNH	Proportion of admissions in past year from nursing home
PADHO	Proportion of admissions in last year from private home
Patient Condition	
BEDRIDDEN	Proportion of patients on census date confined to bed; bed ridden
WALK	Proportion of patients on census date confined to non-ambulatory
DEAF	Proportion of patients on census date confined to deaf
BLIND	Proportion of patients on census date confined to blind
DISTURBED	Proportion of patients on census date confined to disturbed
PSYCH	Proportion of patients on census date confined to admitted from psychiatric hospital
RETARD	Proportion of patients on census date confined to mentally retarded
EPILEPSY1	Proportion of patients on census date confined to epilepsy
EPILEPSY2	Proportion of patients on census date confined to controlled epilepsy
EPILEPSY3	Proportion of patients on census date confined to uncontrolled epilepsy
PALSY	Proportion of patients on census date confined to cerebral palsy
CANCER	Proportion of patients on census date confined to cancer
DIABETES1	Proportion of patients on census date confined to diabetes
DIABETES2	Proportion of patients on census date confined to diabetes, controlled by diet
DIABETES3	Proportion of patients on census date confined to diabetes, controlled by medication
Services	
PT	Physical therapy available; 1=yes; 2=no
OT	Occupational therapy available; 1=yes; 2=no
SPHT	Speech and hearing therapy available; 1=yes; 2=no
COUNSEL	Social service available; 1=yes; 2=no
PNURC	Proportion of patients receiving nursing services
PONEB	Proportion of beds - in one bed rooms
PTHREEB	Proportion of beds - in three bed rooms
PFOURB	Proportion of beds - in four or more bed rooms

Table T-6.1 (Continued)

Case Flow	
ADMISPD	Admission divided by total patient days
Location	
	Health Planning Region
REG1	1 if Connecticut Valley; 0 otherwise
REG2	1 if Central Massachusetts; 0 otherwise
REG3	1 if North West Suburban Boston; 0 otherwise
REG4	1 if North Shore; 0 otherwise
REG5	1 if West South Surburban Boston; 0 otherwise
REG6	1 if South Eastern Massachusetts; 0 otherwise
REG7	1 if Merrimack Valley; 0 otherwise
	Health Systems Area
HSA1	HSA1
HSA2	HSA2
HSA3	HSA3
HSA5	HSA5
HSA6	HSA6
WAGE	County wage rate from (Area Resource file)
Regulatory Variables	
PCPRIV	Proportion census on census date that was private pay
PCMCARE	Proportion census on census date that was Medicare
Licensure and Certification	
CERCARE	= 1 if certified to provide Medicare services; 0 otherwise
PBCARE	Proportion beds certified for Medicare
PBSNF	Proportion beds licensed a Skilled Nursing Facility beds
PBDF	Proportion beds licensed or SNF beds, but not certified as Medicare beds (= PBSNF - PBCARE)
PBRES	Proportion of beds providing rest home care
Facility Characteristics	
NONPROF	= 1 if operated by a nonprofit organization; 0 otherwise
CHAIN	= 1 if operated as part of a group of facilities; 0 otherwise
NCHAIN	Number of facilities operated jointly under one ownership
NCHSQ	NCHAIN squared
MIXED	= 1 if licensed to provide two or more levels of care (SNF - ICF; SNF - RES; ICF - RES; SNF - ICF - RES); 0 otherwise

T-6.5 Empirical Results

T-6.5.1 Functional Form of The Cost Function

Although a multiplicative form for the cost function has theoretical appeal, it has been argued (see Technical Paper #4) that an additive form for the cost function is more amenable to use as the base for rate setting. The current discussion focuses on additive forms for the average cost relationship.[1]

T-6.5.2 The Cost Function Results: Cross-Sectional Analyses

The major objective of the econometric cost function work for this project is to identify variables that vary with operating cost. Causation cannot be uncovered in the single equation cost model, so the results cannot be used to deduce what causes high or low cost behavior. Multivariate regression equations explaining variation in operating cost were estimated using all available variables indicated by previous work, as discussed in Section T-6.3. In general, explanatory power was high for a cross section: the proportion of variation in operating expense is around 66 percent for 1973, 71 percent for 1974, and 72 percent for 1975.

Results with respect to the hypotheses listed in Section T-6.3.2 will be discussed in turn for the ordinary least squares regression of the operating expense for each year on the variables discussed above. The regressions for the three years are reported in Table T-6.2.

T-6.5.3 Economies of Scale

Economists have traditionally been interested in the effect on average cost of scale of production, here represented together by bed capacity, B, and capacity utilization, OCC. Conclusions can be drawn about optimal values for B and OCC only if the average cost function is free to fall and rise as the independent variable increases. The functional forms tried for Massachusetts included (1) quadratic forms for beds and occupancy; (2) grafted

---

[1] A multiplicative relationship, linear in logs, has theoretical appeal, but is felt to be less intuitively acceptable to regulators. Empirically, the log form performed well in explaining cost variation. To compare the explanatory power of this equation with the power of an additive form, it is necessary to compute and square the deviation of its predicted values from the actual values of operating expense, and compare the sum of these squared residuals to the sum of squared residuals in the additive equations fitted for the same data. This computation revealing goodness of fit for the multiplicative equation was performed for 1974 and 1975. The R^2 for 1974 was 0.710 for the converted log linear equation and 0.713 for the best additive form; for 1975, we find $R^2 = 0.725$ for the log linear equation and 0.719 for the best additive form.

Table T-6.2

DETERMINANTS OF AVERAGE OPERATING COSTS:
COMPLETE COST MODELS FOR MASSACHUSETTS
(Standard errors in parentheses)

Explanatory Variable	Regression Coefficients 1973	1974	1975
Constant	34.4131	34.9089	20.8976
Capacity			
BED0	0.0218(.0318)	-0.0243(.0261)	-0.0223(.0303)
BED40	-0.0271*(.0160)	-0.0200(.0139)	-0.0157(.0153)
BED80	0.0069(.0080)	0.0001(.0052)	-0.0029(.0056)
Utilization			
OCC0	-0.2376***(.0369)	-0.2122***(.0598)	-0.1196*(.0648)
0CC90	-0.7890(.6901)	1.2383***(.4469)	0.2005(.3112)
OCC93	0.1872(.7104)	-0.3405(.5943)	-0.4396(.5225)
OCC95	-1.198***(.4413)	-0.5992*(.3126)	0.8581***(.3306)
OCC97	0.4638*(.2722)	0.0952(.2183)	0.2991(.2370)
OCGE93	2.9296*(1.7056)	-1.715(1.0835)	--
OCGE93	--	-1.6967*(1.0021)	--
OCGE95	1.3433(1.0819)	1.2268(.7798)	1.6132*(.8240)
Patient Mix			
PLE49	--	-1.216(3.521)	1.633(4.003)
PSO64	--	4.283*(2.197)	-0.702(2.276)
PLE59	-7.094(2.6296)	--	--
P6569	--	2.766(2.500)	0.804(2.870)
P8089	0.0821(1.4887)	-1.529(1.591)	-0.788(1.621)
POL	7.0736***(2.3696)	3.259*(1.857)	1.625(2.297)
Sex			
PFEM	-9.686(1.0001)	0.605(.769)	1.222(.956)

*Significant at .10 level
**Significant at .05 level
***Significant at .01 level

Table T-6.2 (Cont'd.)

DETERMINANTS OF AVERAGE OPERATING COSTS:
COMPLETE COST MODELS FOR MASSACHUSETTS
(Standard errors in parentheses)

Explanatory Variable	Regression Coefficients 1973	1974	1975
Admission Source			
PADNH	-0.5818(1.3369)	--	-1.994(1.312)
PADHO	2.0395(1.1646)	--	1.222(1.222)
Conditions			
BEDRIDDEN	0.0981(2.3302)	--	--
WALK	--	1.733**(.756)	2.127**(.885)
DEAF	0.3710(3.8219)	-3.658(3.811)	5.599*(3.103)
BLIND	5.8025(6.0472)	-1.166(4.424)	--
DISTURBED	0.0562(.8095)	--	--
PSYCH	--	-1.185(.937)	-0.213(1.102)
RETARD	-2.4350(2.4128)	-2.432(1.716)	1.684(2.669)
EPILEPSY1	-0.6027(5.8503)	--	--
EPILEPSY2	--	-3.122(4.670)	-0.4090(4.623)
EPILEPSY3	--	-4.276(13.990)	2.118(19.273)
PALSY	2.1153(16.4487)	12.200(10.014)	8.413(9.839)
CANCER	--	7.880**(3.130)	4.807(5.833)
DIABETES1	--	2.798(1.969)	--
DIABETES2	--	--	-0.507(2.399)
DIABETES3	--	--	0.947(2.479)
Services			
PT	-0.9597(.7555)	0.861**(.346)	0.850*(.400)
OT	-0.2767(.4797)	--	--
SPHT	0.2518(.4983)	--	--
COUNSEL	0.8684(.4833)	--	--
PNURC	--	0.190(.614)	1.070*(.568)

*Significant at .10 level
**Significant at .05 level
***Significant at .01 level

Table T-6.2 (Cont'd.)

DETERMINANTS OF AVERAGE OPERATING COSTS:
COMPLETE COST MODELS FOR MASSACHUSETTS
(Standard errors in parentheses)

Explanatory Variable	Regression Coefficients 1973	1974	1975
PONEB	--	1.875(1.378)	4.494***(1.448)
PTHREEB	--	-0.333(.548)	-1.028(.627)
PFOURB	--	-1.125*(.613)	-0.98(.711)
Case Flow			
ADMISPD	377.507***(105.553)	262.374***(94.040)	113.801(123.67)
Location			
REG1	-3.4382***(.6605)	-2.674***(.490)	--
REG2	-2.3740***(.6377)	-2.240***(.492)	--
REG3	-1.3853*(.7764)	-0.497(.555)	--
REG4	-0.6099(.6689)	0.400(.472)	--
REG5	-1.8192(.6559)	-0.800*(.484)	--
REG6	-2.6450***(.6190)	-1.161**(.466)	--
REG7	-1.1010(.6988)	-.912*(.541)	--
HSA1	--	--	-1.366**(.645)
HSA2	--	--	-1.937***(.559)
HSA3	--	--	-0.307(.707)
HSA5	--	--	-0.450(.541)
HSA6	--	--	0.803(.538)
Area Descriptors			
WAGE	--	--	3.906(4.819)
Regulation Payment Proportion			
PCPRIV	2.5785***(.8510)	4.548***(.836)	5.154***(.879)

*Significant at .10 level
**Significant at .05 level
***Significant at .01 level

Table T-6.2 (Cont'd.)

DETERMINANTS OF AVERAGE OPERATING COSTS:
COMPLETE COST MODELS FOR MASSACHUSETTS
(Standard errors in parentheses)

Explanatory Variable	Regression Coefficients 1973	1974	1975
Licenses and Certificates			
CERCARE	0.6669(.6532)	--	--
PBCARE	--	3.968***(.870)	3.193***(1.038)
PBSNF	4.2438***(.6986)	--	--
PBDF	--	4.527***(.639)	5.769***(.626)
PBRES	-5.8993**(2.5874)	-5.628***(1.459)	-6.144***(1.764)
Facility Characteristics			
NONPROF	3.5752***(.7841)	3.669***(.665)	2.726***(.581)
CHAIN	--	-.293(.475)	--
NCHAIN	--	--	0.176**(.088)
NCHSQ	--	--	-.00661*(.003)
MIXED	.8036(.4990)	1.089***(.394)	1.331***(.451)
R^2	0.656	0.713	0.719
df	410	419	388
F	18.590	21.644	20.260
S.E.	3.401	2.575	2.821

*Significant at .10 level
**Significant at .05 level
***Significant at .01 level

segments, so that slopes varied for different ranges of the independent variable and the function formed a continuous line; and (3) discontinuous segments, so that segments need not be joined where the ranges meet.[1]

The three alternative forms were tested in the course of searching for an appropriate functional form for the cost function. The three forms were very similar in explanatory power and the coefficients of variables other than the occupancy and scale measures were strikingly similar, both in terms of statistical significance and absolute magnitude. For this reason, and for purposes of consistency with the other state analyses, we focus on the results using the "grafted-segments" approach.[2]

---

[1] Further explanation is in order about why this last form may be appropriate. The functions used by economists to represent short run average cost are generally assumed to be continuous, since variable inputs are generally assumed to be divisible rather than lumpy. This may not be an appropriate functional form for an industry with different regulations for the various ranges of occupancy and bed size. In Massachusetts, input requirements like licensed nursing coverage are specified for units of 40 beds or fewer. An occupancy floor was in effect over the years 1973-1975 imposing a rate penalty for facilities below 90 percent occupancy, and the Rate Setting Commission was considering raising this to 93 percent and 95 percent during the period in question. For this reason, it made sense to test the suitability of discontinuous linear segments in beds and occupancy as a legitimate functional form. Exploring the use of discontinuous linear segments served also as a means of testing in general for nonlinearity.

[2] The results for these alternative forms can be highlighted briefly. The quadratic forms had significant coefficients for beds and beds squared in all three years, and significant coefficients for occupancy and occupancy squared were found in one year. It is interesting to examine the implied minimum average cost values for occupancy and size in beds that emerge from these equations. It is important to realize that these values are not meaningful when the coefficients are insignificant, and should not be used by regulators unless they are significant and stable over the years.

	1973	1974	1975
minimum cost bed size	158	205	238
minimum cost occupancy	93.7%	(121%)	(119%)

As noted in the previous footnote, a discontinuous function may be appropriate when input regulations differ for categories of bed size, and when rates are set differently for facilities above or below an occupancy floor. For this reason, dummies indicating whether the facilities had more than forty beds and more than 80 beds were allowed to enter the equation. They did not enter with coefficients greater than their standard errors with one exception, which did not reach significance at the 10% level. In one equation, 1973, there was an indication that there was a negative impact on average cost for the dummy representing capacity greater than 80 beds.

The annual cross-section equations here differ little in explanatory power. As seen in Table T-6.2, the coefficients of the scale and utilization variables entered do not differ greatly in magnitude or significance. The grafted equation can also be used to show how average cost changes with beds and occupancy. The ranges for these variables were chosen based on regulation. Using 1973 as an example it appears that for beds only the range $40<B<80$ had a significant coefficient; this was negative, indicating returns to higher capacity scale over this range. The graph in Exhibit T-6.1 uses the coefficients that are insignificant as well, and shows a minimum cost bed size somewhere around 80. If the insignificant coefficients are taken to be zero, the equation shows a flat cost curve with respect to beds, falling only over the range 40-80 beds. No bed ranges were significant at the 10% level for 1974 or 1975.

As in the case of capacity, there is an argument for allowing discontinuities in this relationship. Rate regulation in Massachusetts may have an especially strong impact on facilities with occupancy rates less than 90 percent, because of the 90% occupancy floor used for calculating average cost for rate setting purposes. An experiment allowing dummy variables indicating occupancy rates greater than key values to enter the equation showed that the grafted function was overly restrictive: there was a significant increase in the intercept term at 90 percent occupancy in 1973. Another intercept jump occurred at 95 percent, but it was not significant at the 10 percent level. With these two dummies in the equation, the segment between 90 and 93 percent occupancy approached significance (coefficient greater than its standard error) and turned negative; the segment representing occupancy rates greater than or equal to 97 percent became significant, and more positive. Similar results were found for 1974, with a significant decrease in cost for facilities with occupancy greater than 93%, a less than significant decrease occurring at 90%, and an unexpected increase at 95%. For 1975, the coefficient for the dummy showing occupancy over 95% is also positive, and significant at very close to the 5% level. As an example, all coefficients, including the insignificant ones, are used to graph the relationship between occupancy and average cost in 1973 shown in Exhibit T-6.2. When occupancy ranges were used to fit a set of grafted line segments, it was found that all significant coefficients are negative. In all three years, it

Exhibit T-6.1: Bedsize Relationship Estimated from the Complete Model for years 1973, 1974, and 1975

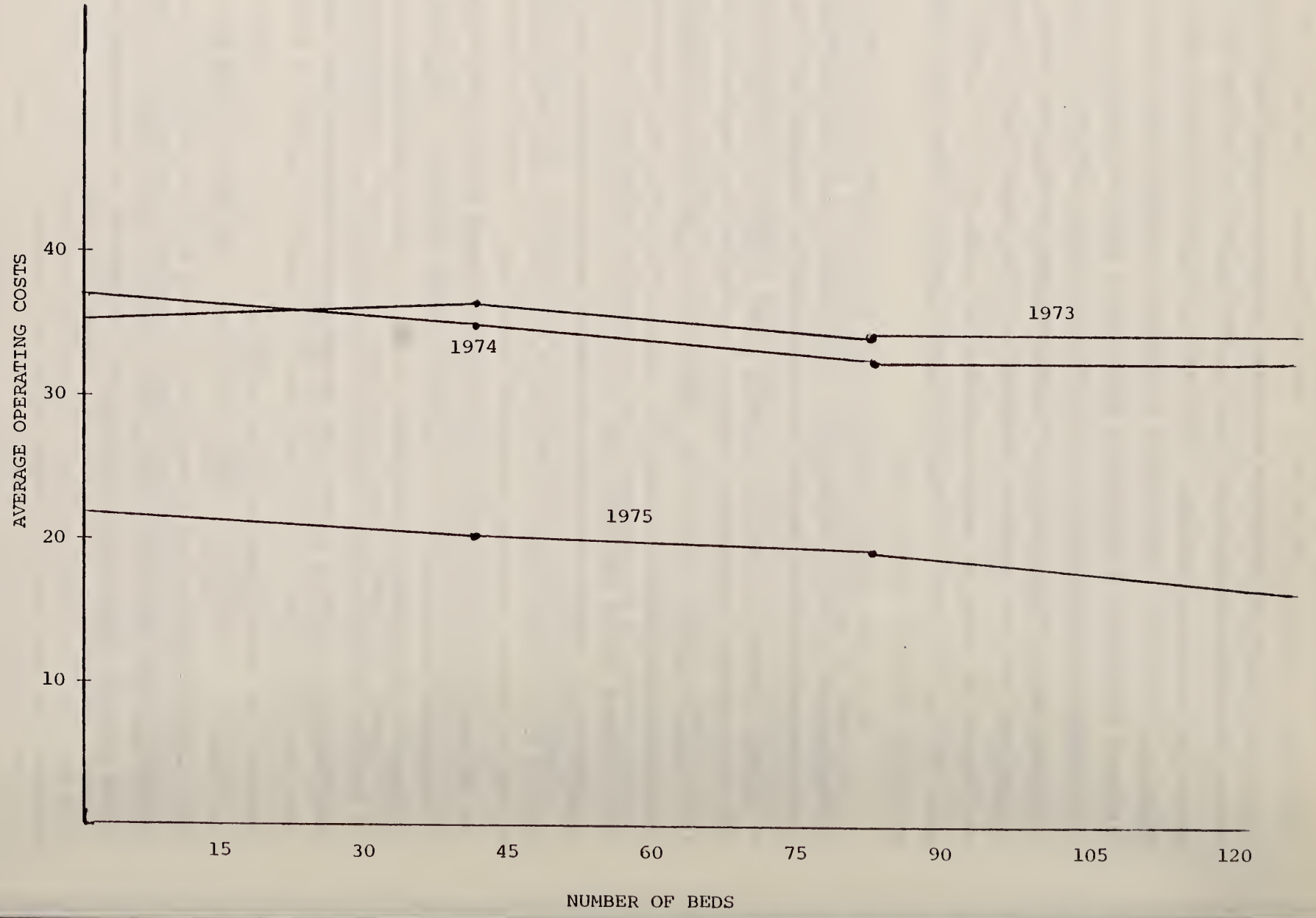
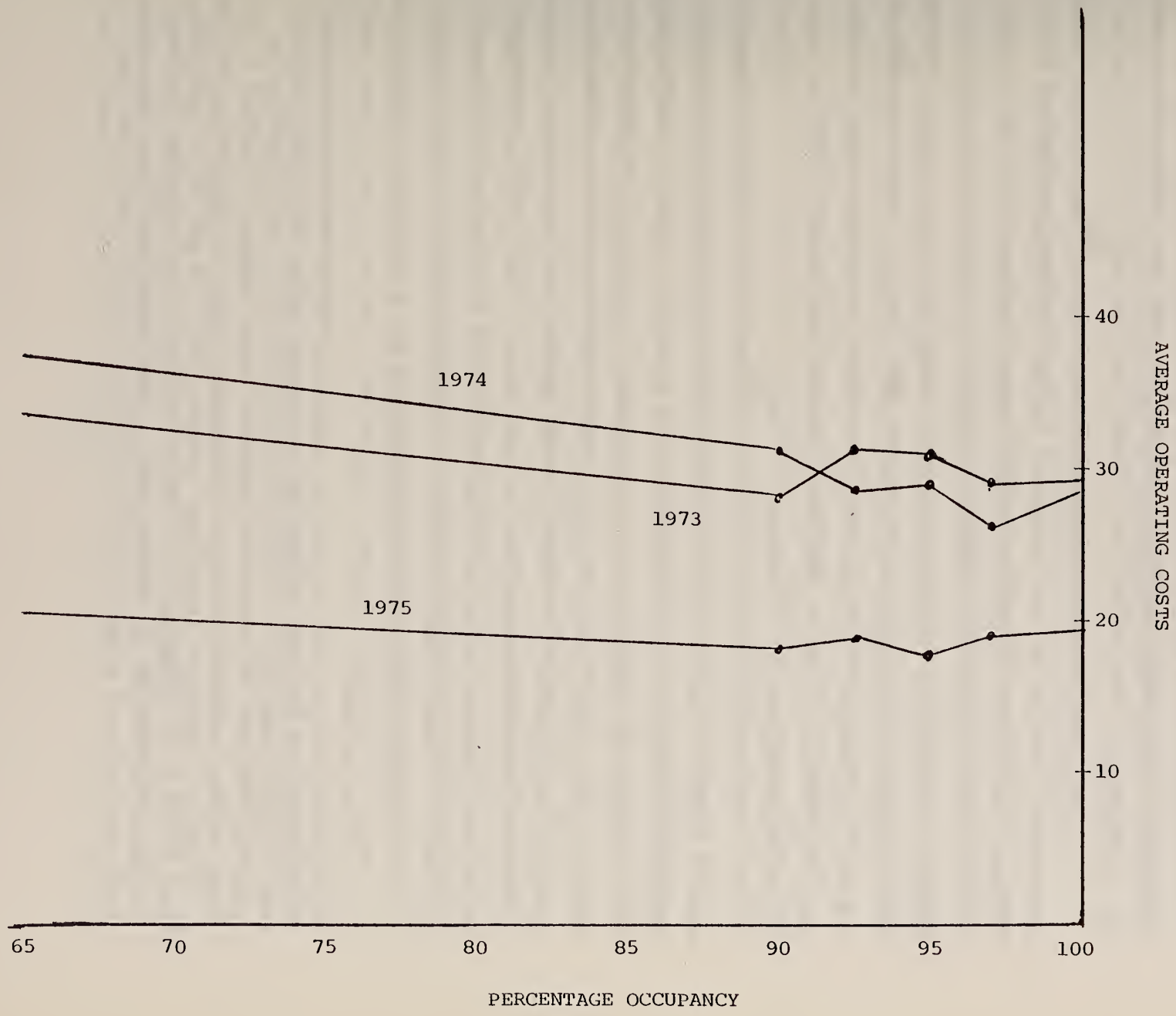


Exhibit T-6.2: Occupancy Relationships Estimated from the Complete Model for years 1973, 1974, and 1975



is clear that average cost decreased with occupancy for capacity utilization less than 90%. In 1973 and 1975, costs also decline significantly between 95 and 97% occupancy.

An examination of the relationship of cost to occupancy showed a minimum average cost near 97% occupancy for both joined and unjoined segments in 1973; near 90% for 1974; and in 1975, the minimum falls near 90% for the continuous relationship, but occurs around 97% when the constraint of continuity is removed.

In general, these experiments with joined and discontinuous segments demonstrate first of all that estimated minimum coverage cost values for occupancy may not be technically stable, since they vary from year to year. Secondly, the experiments illustrate a pitfall of using data which is the product of regulation: the large and significant increases in cost at 90 percent occupancy, and perhaps at 95%, are almost certainly due to regulation and not to production characteristics, so that an estimation that did not take the discontinuities of regulatory effects into account would not model the actual relationship of occupancy to cost.

The conclusions to be drawn with respect to economies or diseconomies of scale or capacity are that these are not very significant, so that regulators should not use these results to press for larger or smaller facility size. Facilities operating at a higher level of capacity utilization do appear to have lower costs per patient day, especially when those below 90% occupancy are compared with others. This implies, not suprisingly, that some elements of operating expense are in fact fixed in the short run and depend on how many patients could be served rather than on actual patient days provided.[1] However, this conclusion is clouded by the fact that occupancy rate is currently used in rate determinations and a facility's decision about

---

[1]Multiplicative equations, fitted by regressing the natural logarithm of average cost on the logs of beds and occupancy (along with other product and facility descriptors) have been used to estimate the proportion of hospital cost that is relatively fixed for minor changes in output. Applying this methodology to the Massachusetts data for operating cost for 1974 and 1975 revealed significant results, with 30% and 32% of operating cost found to be unresponsive to volume changes (i.e., marginal cost was found to be about 70% of average cost). The result for 1973, also significant, was twice as large, with 60% of operating cost found to be fixed.

capacity utilization affected by this: specifically, it can be hypothesized that facilities which are able to attract more private patients (which are more expensive, as indicated by payment proportions to be discussed below) are less concerned with meeting the Massachusetts Rate Setting Commission's occupancy floor.

We now turn to other output characteristics of nursing homes. The type of product that the nursing home produces is expected to affect average cost. Both the service intensity and the patient type served may be seen as product dimensions. We first examine service intensity.

T-6.5.4 Services Intensity.

In the Massachusetts case, the proportion of beds that are licensed to provide various levels of care, and the proportion of beds certified to provide services under Medicare, are the best available indicators of the type of service provided. In addition, variables for the availability of certain services and for the use of nursing services are present in some years.

Before the coefficients of the service intensity variables are discussed, it is important to consider whether the levels of care (SNF, ICF, and mixed SNF, ICF, and/or personal care beds) are so different that facilities operate on different production functions and thus on different cost functions. A test for this is to fit regressions for the three different level groups using the same variables and to test whether the coefficients differ among these regressions compared with the regression on the full population. Because there are relatively few freestanding SNFs, it was necessary to reduce the variable list to conserve degrees of freedom. The four resulting regressions are compared in tabular form in Appendix T-6.3. It is clear that the coefficients are in fact significantly different, and Chow tests confirm this. This throws into question analysis that fits a single cost function for both freestanding facility types and for mixed facilities for Massachusetts.

In spite of these doubts, it is possible to examine the impact of type of licensure on cost in a continuous fashion, using the proportion of beds at each licensure level. Proportion variables are again useful, since if beds are fully occupied, the coefficient of the proportion variable in the linear form shows how much an additional bed costs to operate for a day. With ICF beds as the left-out proportion, we find for 1973 a $4.24 increment

for SNF beds over ICF beds and $5.89 decrement for rest home beds. These coefficients have the expected signs, since SNF case is more service intensive than ICF or personal care. The Medicare certification dummy variable adds $0.66 to average cost.

In 1974 and 1975, SNF beds can be further broken down into those that are certified to provide Medicare services (PBCARE) and those that are not (PBDF). Certification of beds as SNF beds added $3.96 for Medicare beds and $4.52 for non-Medicare beds in 1974; the figures for 1975 are $3.19 and $5.77. At first glance it is surprising that non-Medicare SNF beds cost more. However, these equations include patient proportion by payment type, so that the cost of a Medicare bed is in a sense being broken down into the cost of caring for a Medicare certified patient and the cost of maintaining a Medicare bed, whether filled by a Medicare, Medicaid or private patient.[1] Medicare beds are more costly to maintain no matter what type of patient is occupying them. Personal care (rest home) beds were on average $5.88 less costly than ICF beds in all years according to the coefficients of PBRES. These coefficients appear quite stable in magnitude over the three years, during a time when licensure standards did not change. The increases over the three years for the more service intensive levels of care can be explained as the result of inflation.

No service availability variables are significant for 1973. For 1974 and 1975, only physical therapy availability is reported. Although this was not significant in 1973, it is for these later years, adding $.86 and $.85 to average cost, respectively. The proportion of patients receiving nursing services has a significant impact on cost only in 1975; the coefficient shows an increase of $1.07 for each patient requiring nursing services.

The general service intensity hypothesis, that more intensive services cost more, is consistent with these results. However, it is clear that much better data than Massachusetts collects would be necessary to set prices related to particular dimensions of service provision.

---

[1]Before these proportions are added, the coefficients of PBCARE and PBDF were $7.30 and $5.45 for 1974, and $7.93 and $6.42 for 1975 respectively.

T-6.5.5 Patient Mix Variables

The variables available to describe patient mix for Massachusetts nursing homes are very limited in that they refer only to diagnosis, age, sex, and admission source. Measures of independence in activities of daily living, believed by many to predict nursing and personal care needs of patients, are not available. Although some diagnosis categories are not present in all three years, and the age ranges are not exactly the same for all years, it is interesting to review the impact of these variables on operating expense over the three years considered here. Few of these variables are significant, their coefficients change in significance, sign, and magnitude from year to year, and as a group they do not add a great deal to the explanatory power of the cost equation.

For 1973, there are no diagnosis categories with significant coefficents. The only significant age proportion is that for patients 90 years of age or greater; these patients are estimated to cost $7.07 more per day than the left-out category, patients 60 to 79.

In 1974, more diagnosis categories were reported. Each bedridden patient was found to add $1.73 to total cost, and each cancer patient $7.88. Again, the proportion 90 years of age or older had a positive significant coefficient, adding $3.26 to cost for each patient. In addition, patients in the 50 to 64 age group added a significant amount, $4.28 over the null case (70 to 80).

In 1975, the proportion with cancer is insignificant. The coefficient for the proportion bedridden is again positive and significant, and rises to $2.13 per patient. The coefficient on the proportion deaf is for the first time highly significant and positive, showing an addition of $5.60 to cost for deaf patients. Since there is little in the literature to suggest that deafness is a condition which adds significantly to cost, we believe this result to be an artifact. This demonstrates how any such patient indicator may take a high value in one year in particularly high cost

T-6-25

facilities, and thus appear to raise cost: causality should not be inferred.[1]

The proportion of patients who are female entered positively but was never significant.

These results on patient mix will be disappointing to those who believe that it is patient mix differences that call forth different levels resources, and hence differing costs, across nursing homes: even though the data are not ideal, there is little indication that case mix makes much difference to cost. The level at which beds in the facility are staffed (SNF, IC, or RES) and, as will be shown below, the method of payment, have much more power to explain cost per patient day than do patient characteristics, at least as available in the Massachusetts data.

T-6.5.6 Case Flow

The coefficient for the variable ADMISPD shows the addition to total cost of an additional admission, all else constant.[1]

This variable was significant for 1973 and 1974, but not for 1975. The estimated cost per admission was $377.51 in 1973 and $262.37 in 1974. Proportion of discharged patients in various length of stay categories did not work as well as ADMISPD, as judged by comparing standard errors of estimate for equations using each formulation.

The results with respect to ADMISPD confirm that a higher case flow (shorter length of stay) is associated with higher average per diem cost.

---

[1] It is interesting that some patient proportion variables that were not significant in the additive form achieved significance in log linear forms fit for the same data. For example, in 1973 it was found that mentally retarded patients were less costly, with a coefficient of -.304 significant at the 5% level. The coefficient of proportion variables in the log linear form can be interpreted as a percentage impact: if all patients were mentally retarded, average costs would be multiplied by $e^{-.304} = .74$, implying that these patients are 26% less costly than those without this condition. Cancer patients in 1974 were 46% more costly than those without cancer, and bedridden patients were 9% more expensive than ambulatory patients in 1975. These results are not overwhelmingly different from the additive results, but they do fit with the concept of proportional resource allocation by patient type implicit in Winn and McCaffree's analysis.

[1] $AC = b_o + b_1 \text{ ADMISPD} + b_i x_i$ where $\text{ADMISPD} = \frac{\text{admission}}{PD}$

$TC = PD\ AC$

$= PD\ b_o + b_1 \text{ admission} + PD\ b_i x_i$

Therefore b_1 is the addition to total cost of an additional admission, all else (including patient days) constant.

T-6.5.7 Input Prices

The general indicators of input price levels, available for counties for 1975 from the Area Resource File, were not significant whether included with or without the other input price proxies, regional variables. Counties in Massachusetts are much less compact than counties in other parts of the country, and cannot be assumed to describe labor market areas even in a general way.

The health planning areas and Health Systems Agency areas are not at all congruent with counties and are arranged to take in more unified market areas. In 1973, average costs were significantly less than costs in the left-out category, the Boston area, for all areas but the North Shore. The rural western areas and the southeastern area which includes Cape Cod exhibited average operating costs that were less than Boston on average; and three of the four more urban areas near Boston had average costs between one and two dollars less than costs in the Boston health planning area. In 1974, the differences from Boston are reduced for all regions, but the ordering of the regions in cost remains the same. The new Health Systems Areas are apparently not as effective in capturing market differences, since only the western areas have significant coefficients in 1975. In sum, it appears that there are significant regional differences in costs, which may be growing smaller over time. The growing potential for union organization of nursing home workers across the Commonwealth may be playing a part in this trend.

T-6.5.8 Regulation

The major regulation variables within the state are the proportions of public, private, and Medicare patients served, which it is hoped will capture the incentive a facility has to respond to rate regulation. (All facilities must respond to licensure regulation.) In the years that the proportion of Medicare patients served is available, this was found to have a large positive impact on cost: it appears that the cost per percentage increment in the number of each Medicare patient added $0.14 per day in 1973 over the cost of a public patient, $0.12 in 1974, and $0.29 in 1975. The other variables held constant do include proportion of Medicare certified beds in 1974 and 1975, and cross-year comparisons are not completely legitimate because of changes in the variable set, but the difference in the magnitude of the coefficients is puzzling. The sign is as expected, since Medicare reimbursment covers a higher proportion of average cost than does reimbursement for public patients; in addition, of course, these patients are found in facilities certified to provide intensive rehabilitation services, so the patient variable is also capturing a product dimension.

The proportion of patients who are private pay also had positive and highly significant coefficients, as expected. The coefficients implied that facilities that were 100% private pay would cost $2.57 per patient day more in 1973, $4.55 more in 1974, and $5.15 more in 1975 than facilities serving all public patients, all else constant. This can be interpreted to mean that facilities serving more public patients must be more cost-conscious, and can produce similar services for less; or, as noted above, a higher private pay patient load may be associated with aspects of nursing home service, for example amenities or higher quality care, that are not captured by other variables.

T-6.5.9 Additional Facility Characteristics

Nonprofit ownership adds significantly to cost in every year. The coefficients show additions of $3.57 in 1973, $3.67 in 1974 and $2.73 in 1975. Even holding other aspects of services, location, and patient popuation constant, nonprofit facilities cost more.

Nursing home services produced in mixed facilities, where more than one licensed level of care is provided, are also more expensive. The differential is $0.80 per patient day in 1973, $1.09 in 1974, and $1.33 in 1975, all significant at better than the 5% level.

Chain ownership tended to reduce cost in 1974, but the coefficient was not significant. In 1975, the number of facilities which were jointly operated was reported on the Facility Survey. Use of this variable and its square provide an interesting result: costs increase (at a decreasing rate) as the number of facilities in the chain increases, with average cost reaching a maximum at around 13 facilities. Since a small percentage of the facilities were operated jointly in groups of 13 or more, most nursing homes in "chains" in Massachusetts are apparently not reaping economies of joint ownership.

The services of capital are captured by the proportion of beds in one, two, three, and four or more bed (ward) rooms. It was significantly less expensive by $1.13 per patient day to serve a patient in a ward bed in 1974; the coefficients for the proportion of beds in one and three bed rooms were insignificant but bore the expected relationship to the left-out proportion, beds in two bed rooms (+$1.88 and -$0.33 respectively). In 1975, the proportion of beds in single rooms took on a large and highly significant coefficient,

indicating single rooms were $5.49 more expensive to operate than beds in two bed rooms. This coefficient is different enough from the preceding year to raise a question about whether it is an artifact, due perhaps to the association of single rooms with newer facilities providing more sophisticated services to a private-pay population that demands other special services along with the privacy of single rooms. Again, the other two variables took on coefficients of expected relative magnitude, (-$1.03 and -$0.98) but they were insignificant.

As in the New York analysis, beta coefficients were examined in a further effort to assess the impact of different variables on costs. For various groups of variables, the beta coefficients were summed.[1] The table below reports the summed beta values for 1975 using 12 groups of variables.

Table T-6.3

Summed Beta Coefficients

Variable Group	Beta Coefficient Sum
Services Offered	0.0707
Services Provided	0.0571
Capital Sources	0.0372
Input Prices	0.0508
Urban-Rural	0.0304
Admission/Patient day	0.5540
Patient Age	0.0153
Patient Diagnosis	0.0264
Activities of Daily Living	0.0822
Impairments	-0.0517
Admission Source	-0.0090

[1] The beta coefficients are simply normalized regression coefficients and bear a straightforward relationship to the unnormalized regression coefficients, being equal to the unnormalized coefficient times the ratio of standard deviations for the independent (S_x) and dependent (S_y) variables:

$$b_x \cdot \frac{S_x}{S_y}$$

A beta coefficient of, say, .25 can be interpreted to mean that a one-standard deviation change in the independent variable leads to a .25 standard deviation change in average operating costs, the dependent variable. (It is interesting to note that the beta coefficient for the independent variable in a bivariate regression model is identically equal to the simple correlation between the two variables.)

Using the beta coefficient sums as measures of the impact of different groups of variables we find that case flow as measured by the number of admissions per patient day has the largest impact on costs; a one standard deviation increase in the variable leads to a 0.55 standard deviation increase in per diem operating costs. Other groups of variables reported in the table do not even come close to the case flow variable in terms of the magnitude of impact. The measures of patients' activities of daily living (ADL) have the second largest impact on costs. A one standard deviation increase in each ADL variable leads to a net 0.082 standard deviation increase in costs, ceteris paribus.

Groups of Variables

The models fit here have a relatively high degree of multi-collinearity due to the large number of independent variables, many of which measure similar factors. In Table T-6.4 we report on attempts to assess the statistical significance of groups of variables. The F-tests applied are not as sensitive to multi-collinearity problems as the individual t-tests used elsewhere in this paper. The variable groups report in the table are 1975 data. The variable groups are significant at the 0.05 level or better if the F-statistic is greater than or equal to the criterion value. Summarizing, the following groups of variables are significant at the 0.05 level or better: (1) service intensity, (2) patients' payment source, and (3) additional facility characteristics. The three groups of patient mix descriptors which were tried were insignificant. The measures of scale and utilization were also insignificant. The reader should bear in mind that these results depend on the specification of the model within which the variable groups have been tested.

T-6.6 Components of Cost

Relationships were estimated using five different components of average cost per patient day as dependent variables. These were: nursing expense, dietary expense, maintenance expense, laundry expense, and total

Table T-6.4

Residual F-Test Results

	Variable Group	F-Statistic	Significance Criterion for 0.05 level
(1)	Occupancy (OCC0-OCC95)	1.79	2.12
(2)	Beds (BED0-BED80)	1.24	2.64
(3)	Occupancy and Beds	1.73	1.90
(4)	Patients' Diagnoses (DEAF, RETARD, EPILEPSY1, EPILEPSY2, EPILEPSY3, PALSY, CANCER, DIABETES1, DIABETES2, DIABETES3)	1.14	1.85
(5)	Patients' Age and Sex (PLE49, P5064, P6569, P8089, POLD, PFEM)	1.80	3.02
(6)	Activities of Daily Living, Diagnoses, Age, and Sex (WALK, plus variables listed in (4) and (5).	1.26	1.83
(7)	Service Intensity (PT, PNURC, PONEB, PTHREEB, PFOURB)	3.85	3.02
(8)	Payment Sources (PCPRIV, PCMCARE)	22.17	3.02
(9)	Facility Characteristics (NONPROF, NCHAIN, NCHSQ, MIXED)	7.91	3.83

salary expense per patient day. Scale and utilization variables were used in every equation, but others were allowed to enter only if their coefficients were greater than their standard errors. The proportion of variance explained in the various models is reviewed in the chart below:

	1973	1974	1975
Nursing	0.678	0.668	0.702
Dietary	0.540	0.507	0.499
Maintenance	0.419	0.389	0.389
Laundry	0.314	0.188	0.267
Salary	0.604	0.67	0.70

Nursing expense was expected to be related to level of care and patient characteristics. Dietary costs were expected to exhibit some economies of scale because of the fixed cost of operating a kitchen no matter what the size of the facility or the number of patients served. Maintenance costs per patient day were expected to respond strongly to utilization since capital facilities must be maintained whether beds are filled or not; this component of operating cost is hypothesized to be more fixed than some others. Laundry cost, on the other hand (like the raw food portion of dietary expense) does depend on the daily census: facilities can contract for laundry services in a way that is responsive to changes in utilization, and there need not be fixed costs. Thus neither economies of scale nor of utilization are expected. The salary expense equation is expected to look very like the equation explaining OPEXP, since salaries are a very high proportion of operating expense. Nonprofit homes and homes serving more private patients are more costly. Economies of delivering meals may be shown by the fact that costs are significantly higher when more beds are in one bed room, and tend to be lower when more beds are in wards, in contrast to two- or three-bed rooms. Only one age group, 65 to 69, entered the 1975 equation; the positive sign on this proportion variable may be associated with more demands for better food from this younger and presumably healthier age group. Diabetic and cancer patients, who have special diet problems, were found to cost more. It is difficult to explain the high significant coefficient for the proportion of deaf patients except as an artifact. Economies made possible by joint purchasing or joint contracts for food service apparently do not influence diet costs until a scale of 10 facilities is reached, according to the chain variables.

T-6.7 The Cost Function Over Time

T-6.7.1 Introduction and Hypotheses

In most industries it is possible to assume that a cost relationship fitted in one year will hold in subsequent years. Changing input prices are accounted for in the ideal cost function, so that cost charges due to changing relative prices and overall inflation should theoretically be accounted for. The underlying production relationship is expected to be stable, and producers move along the production frontier in response to changing prices. In a general sense, this is corroborated by the Massachusetts results for 1973, 1974, and 1975. Scale and product differences do have reasonably stable impacts on average cost. After inflation has been adjusted for, one would expect that the best 1973 equation would predict a substantial amount of the variation in 1974 and 1975 actual costs. This is exactly the stability that is needed if a cost function fitted for one period is to form the basis for a rate setting system to be applied in a later period.

However, while relationships have been found for each year between the dependent variable, average operating cost, and scale, product, and facility variables, it would be a mistake to accept these as causal impacts without further investigation. Approaching the cost relationship through changes in average cost can cast a different light on the short run impact on cost of changes in independent variables, and can incidentally provide insight into some other questions of interest to policy makers. This work is exploratory in nature, so the initial hypotheses are stated in very general terms:

H.1 Independent variables which are found in cross section to be significantly associated with average cost will also be significant in a first difference form relating changes in cost to changes in the independent variables. These variables include: occupancy rate, output descriptors, and patient payment mechanisms.

H.2 Certain product types have been more subject to inflation than others, so that facilities producing these outputs in several years (i.e., little change in product type) will find their average per diem operating cost rising more than the average.

H.3 Certain types of facilities are better than others at controlling cost.

H.4 Facilties serving more public patients are more subject to rate regulation, and have strong incentives to control costs.

H.5 Facilities with cost components above announced rate ceilings will have a strong incentive to control cost.

H.6 Facilities operating below cost component ceilings will face less incentive to keep costs down, and may even be expected to increase cost toward the announced ceiling.

T-6.7.2 Empirical Results

The hypotheses are tested by using two changes in operating expense for individual facilities. Certain variables were selected as potentially important in a first-difference form because of their impact in cross section and because they did change (in contrast to bed size, ownership, location, and other variables which did not change from year to year). Other variables describing certain facility characteristics which were constant from year to year were either included a priori (ownership, chain status) or were allowed to enter if their coefficients were greater than their standard errors. Product type descriptors (patient mix, services) were allowed to enter on the same basis. The proportion of patients who were private pay in the base year was used to indicate rate regulation pressure on the facility. Finally, it was possible to estimate what ceiling applied to "variable cost" (as defined by Massachusetts rate regulations) and to nursing salaries per patient day for each facility in 1973. (This ceiling was actualy applied to the facility's reported cost for 1971, which was inflated to find the 1973 interim rate. The hypothesized impact of the announcement of the ceiling will be presented in more detail below). Underruns and overruns with respect to these estimated ceilings were used as independent variables to gauge their impact on operating cost one and two years later.

Results are presented in Table T-6.5. Almost half the variation in ΔOPEXP is explained for 1973-74 (R^2= 0.480).

The five change variables, ΔOCC, ΔPOLD, ΔADMSPD, ΔPPRIV, and ΔPCMCARE, were selected because they were utilization and product descriptors that were significant in the earlier cross section analyses. Changes in the occupancy rate were indeed associated with average cost changes as expected: costs rose 8.5% for each percentage point drop in average occupancy over the first period, and rose 6.1% per point decrease in the second period. It is interesting to compare this year-to-year response rate with the cross-sectional impact of occupancy (when used as a continuous variable).

	Cross-Sectional Impact: OCC	Year to Year Impact: ΔOCC
1973	-.2702***	—
1973-1974	--	-.0852***
1974	-.1171***	—
1974-1975	—	-.0611***
1975	-.0674*	--

All but the coefficient for occupancy in the 1973 cross-section equation are within 5% of each other. This lends additional support to these coefficients.

The changes in proportion of patients over 90 years of age and the change in admission rate per patient day were not significant in the first-difference equation.

For the first-difference between 1973-1974, the changes in the proportions of private and Medicare patients were significant in explaining cost change. The coefficients for the change in private pay patients were between 12% and 28% of the cross-section coefficients for this variable.

Table T-6. 5

FIRST DIFFERENCE COST MODELS, MASSACHUSETTS - 1973-1974 and 1975-1976
(Standard errors in parentheses)

Independent Variable	1973-1974	1975-1976
Constant	2.1440	1.8868
First Difference Variables		
ΔOCC	-0.0851***(0.0149)	-0.0611**(0.0233)
ΔPOCD	-0.7806(0.7094)	-0.2365(0.9692)
ΔADPD	-55.4318(61.1929)	59.5206(81.6193)
ΔPRIV	0.7926*(0.4427)	0.6279(0.9254)
ΔPCARE	6.9949***(1.6607)	0.9447(5.5458)
ΔBSNF	1.5577***(0.5827)	--
ΔBRES	-28.6736(23.6035)	-3.2257*(1.4972)
ΔBCARE	--	-0.1351(0.5380)
ΔBDF	--	1.0619(0.7678)
Ceiling-Related Variables		
SALARY-UNDER	-0.0492(0.0580)	0.0970(0.0733)
COST-UNDER	-0.1132(0.0818)	0.1209(0.1073)
SALARY-OVER	-0.3922*(0.1449)	-0.0182(0.1881)
COST-OVER	-0.0087(0.0434)	-0.0116(0.0566)
Capacity		
BED0	0.0107(0.0092)	--
BED40	--	--
BED80	--	--
BGT80	--	0.2176(0.1987)

*Significant at 0.10 level
**Significant at 0.05 level
***Significant at 0.01 level

Table T-6.5 (Cont'd.)

FIRST DIFFERENCE COST MODELS, MASSACHUSETTS - 1973-1974 and 1975-1976
(Standard errors in parentheses)

Independent Variable	1973-1974	1975-1976
Utilization		
OCC0	--	--
OCC90	-0.1654(0.1644)	--
OCC93	-0.4463***(0.1494)	--
OCC95	--	--
OCC97	--	--
OCGE90	--	--
OCGE93	0.6946(0.4856)	0.4939**(0.2060)
OCGE95	--	--
OCGE97	0.2756**(0.1249)	--
Patient Mix		
PFEM	-0.5030(0.3172)	-0.5628(0.4086)
PADNH	--	--
PADHO	0.4490(0.3423)	--
EPILEPSY3	--	18.9798**(8.4542)
PALSY	--	11.4879*(5.4952)
RETARD	--	-0.8975(0.7375)
CANCER	--	-2.9352(1.7375)
Services		
PT	0.3188(0.2102)	--
OT	--	--
SPHT	--	--
COUNSEL	-0.2016(0.1513)	--

*Significant at 0.10 level
**Significant at 0.05 level
***Significant at 0.01 level

Table T-6.5 (Cont'd.)

FIRST DIFFERENCE COST MODELS, MASSACHUSETTS - 1973-1974 and 1975-1976
(Standard errors in parentheses)

Independent Variable	1973-1974	1975-1976
PNURC	--	-0.5743(0.3380)
PONEB(1974)	0.8811(0.6412)	--
PTHREEB(1974)	--	--
PFOURB(1974)	-0.2854(0.2252)	0.4367(0.2957)
Case Flow		
ADMISPD	--	136.8040***(33.6275)
Location		
REGCEN3	-0.3021*(0.1611)	--
REGCVA4	--	0.4092*(0.2268)
REGMUA	--	-0.5102*(0.2656)
REGNSH4	--	-0.2733(0.1964)
Area Descriptors		
WAGE	--	--
Regulation		
Payment Proportion		
PCPRIV	0.7789**(0.3387)	1.8542***(0.4137)
Licenses and Certificates		
CERCARE	6.3566***(1.1567)	--
PBCARE	--	-1.1410**(0.4755)
PBSNF	0.5031**(0.2166)	--

*Significant at 0.10 level
**Significant at 0.05 level
***Significant at 0.01 level

Table T-6.5 (Cont'd.)

FIRST DIFFERENCE COST MODELS, MASSACHUSETTS - 1973-1974 and 1975-1976
(Standard errors in parentheses)

Independent Variable	1973-1974	1975-1976
PBDF	--	--
PBRES	--	-1.6950(1.2984)
Facility Characteristics		
NONPROF	0.0458(0.2484)	0.4641(0.2946)
CHAIN(1974)	-0.4874**(0.1927)	-0.2171(0.2544)
NCHAIN	--	--
NCHSQ	--	--
MIXED	--	0.4140**(0.1823)
R^2	0.4805	0.355
Adjusted R^2	0.4209	0.272
Sample Size	273	273
Standard Error	0.8239	1.063

*Significant at 0.10 level
**Significant at 0.05 level
***Significant at 0.01 level

	Cross-Sectional Impact: PPRIV	Year to Year Impact: ΔPPRIV
1973	2.857***	--
1973-1974	--	.793*
1974	4.548***	--
1974-1975	--	.628
1975	5.154***	--

The smaller magnitude (and lower significance) lends support to one interpretation of the proportion private pay variable in cross-section. The cross-section results indicate that a private patient added $2.86, $4.55, and $5.15 to average cost in each year; however, a change in proportion private pay is indicating a market reorientation of facilities, which does not change much for small year to year changes in the actual number of private patients served. In fact, this variable has no significant impact on cost changes in 1974-75, even though in the cross-sectional models it is highly significant.

The change in the proportion of Medicare patients served was extremely small, and was zero for many facilities. Its coefficient was also much smaller than the coefficient for the comparable cross-section variables indicating again that the cross-section coeffficents cannot be relied upon to show responses of average cost to changes in this aspect of patient mix.

Change in the proportion of SNF beds had a significant impact on cost in 1973 to 1974, but this was smaller than the expected cross-section coefficient. The mean change was negative, and this was at time in which nursing home operators are known to have down-graded their beds to avoid the higher SNF standard. The coefficient implies that the cost savings due to providing a lower level of care are not as great as would be predicted in cross-section: the facility may relicense some of its SNF beds as ICF beds, but still maintain back-up SNF services and/or a SNF orientation toward intensive treatment and rehabilitation. Similarly, the savings due to increasing the proportion of rest home beds predicted in cross-section is larger than that indicated by the first difference equation, when significant.

Base year variables can be used to explain changes in cost if it is believed that the output type does not change much along these dimensions from year to year, so the coefficients on the level variables shows differential impact of inflation on these product aspects. (This is corroborated by the correlation of available patient descriptors across years). For example, the coefficient of $PCMCARE_{1975}$ with $\Delta OPEXP_{1973-74}$ as the dependent variable is significant at the 1% level and large, 6.356. It can be interpreted as showing that facilities serving 100% Medicare patients (PCMCARE=1) would find average cost rising $6.35 between 1973 and 1974 over the cost increase of facilities serving only public patients; those with proportions of Medicare patients between 0 and 1 would experience proportionate increases. Facilities maintaining a high proportion of licensed SNF beds also experienced higher cost increases; those maintaining more beds in four or more bed rooms had lower than average increases 1973 to 1974 and higher than average 1974 to 1975.[1] Facilities providing physical therapy in 1973 had higher cost increases and those providing social services had lower cost increases.

Homes providing nursing services to more of their patients in 1974 (PNURC) had significantly lower increases in average cost between 1974 and 1975. For the most part (with the exception of the last variable which requires more discussion) these coefficients indicate, not surprisingly, that intensive services, requiring more inputs, will increase in absolute dollar cost more than will less intensive services: inflation tends to have a multiplicative effect on base costs.

The negative coefficient for PNURC between 1974 and 1975 is an exception to this conclusion, and is consistent with a particular impact of regulatory change that occurred for that year. In 1974, nursing cost ceilings which had previously applied to <u>total</u> nursing salaries were altered so that they applied only to licensed nursing salaries, nurse aide salaries,

---

[1] These coefficients are not significant at the 0.10 level although they are greater than their standard errors.

were included, along with most other components of operating expense, in what is called in Massachusetts "variable cost," subject to the variable cost ceiling. After the industry protested that this was dissuading providers from employing the most efficient combination of licensed and unlicensed personnel, and encouraging the over-employment of licensed nursing personnel, the regulation on ceilings was changed back again in 1975. Facilities providing more nursing services to patients might be expected to experience this effect most strongly, as an incentive to control cost. The increased opportunity they had to substitute aides for LPNs and RNs in 1975 may account for the negative relationship between PNURC and cost increases.

With respect to patient characteristics, it is notable that no base year diagnosis or age proportions had coefficients greater than their standard error for the 1973-74 first difference equation. For 1974-75, the proportions of patients with uncontrolled epilepsy and cerebral palsy appeared to add significant amounts to average cost, but these variables are so small that they may be picking up the effect of specialization for a few costly facilities. The relative cost of treating cancer patients apparently fell between 1974 and 1975 (the coefficient of $CANCER_{74}$ is -2.935 and is significant at the 10% level).

No age proportions were significant for either year. Surprisingly, the proportion of female patients, which was not important in explaining levels of average cost, had coefficients greater than their standard errors for both first difference equations; a nursing home caring only for women would have had about $0.50 lower than average cost increase between 1973 and 1974, and about $0.56 lower cost increase from 1974 to 1975. Facilities with a high rate of admission per patient day experienced significantly higher cost increases between 1974 and 1975.

What types of facilities were best able to keep costs down? Apparently nonprofit status did not make a significant difference to cost changes. Operation as a chain was significant in reducing cost increases 1973 to 1974 by about $0.49; the coefficient for 1974-75 was also negative, but not significant. Mixed facilities, those providing more than one licensed level of care, had significantly higher cost increases for 1974-75 by about $0.41. Facilities located in several regions also found average costs changing at a different rate from Boston changes.

Bed capacity variables were not significant in explaining cost changes, although there is weak evidence that the largest facilities (beds greater than 80) had higher cost increases. High occupancy in the base year led to higher than average cost increases, however. This may simply be because a high occupancy facility is more likey to have lower occupancy the next year, since there is more room for utilization to move down than up; but it may also be because more costs are covered by the Medicaid rate in a high occupancy facility, so that there is less pressure for cost control.

Finally it is interesting to look at the estimated coefficients for the ceiling variables. In Massachusetts an interim rate is paid for care provided in the rate year. The interim rate for 1973 was based on a facility's actual cost for 1971, screened by ceilings computed from industry data again from 1971. The interim payment is subject to adjustment when actual costs are known, and the final rate for 1973 was set (in 1974 or 1975) using actual 1973 costs screened by ceiling based on all nursing home costs for 1973. A facility therefore cannot know its final rate during the rate year, or even what ceilings its costs will be subject to. Management does know the 1971 ceilings used to set the interim rate, however, and for cash flow reasons alone, facilities have an incentive to operate within their interim rate. It is somewhat of a leap to assume that operators compared current 1973 costs with the 1971 ceilings (inflated) to get some idea of what they would do with respect to the final 1973 ceilings. The deviation of 1973 nursing salary expense per patient day from the ceiling appropriate to each facility was computed, with SALARY-OVER representing excess over the ceiling (always positive or zero) and SALARY-UNDER representing underruns below the ceiling (always negative or zero). In like manner, COST-OVER and COST-UNDER are the positive or negative deviations from the variable cost ceilings for each nursing home. The coefficients for the change in average cost from 1973 to 1974 all have the expected negative sign: facilities above the announced ceilings hold cost increases down, while those below the ceilings experience greater than average increases.

However, the results are disappointing in that only the coefficient for SALARY-OVER is significant. Results might be improved by looking at changes in the actual regulated components (nursing salaries, variable expense) rather than total operating expense per patient day. This might allow more specific testing of the pressure implicit in rate ceilings for

lower facility costs to move up toward a ceiling. The indication in the current test is that this pressure is weaker than the downward pressure on facilities above the nursing ceiling to cut costs: the coefficient on SALARY-OVER shows $0.39 lower cost increase for each dollar above the nursing ceiling, while SALARY-UNDER shows a $0.05 increase for each dollar below it. On the other hand, costs above the variable cost ceiling (COST-OVER) induce little or no downward movement (an insignificant $0.01 per dollar excess) while the coefficient of COST-UNDER, greater than its standard error, shows an $0.11 increase for every dollar below the ceiling. The 1973 variables were also used in the 1974-75 equations, to see if cost control impacts of rate ceilings might take longer than a year to work out. They were not particularly helpful in explaining costs, and underruns in 1973 appear to induce <u>decreases</u> in cost in 1974-75, which may merely mean that lower cost facilities (standardizing for licensed level) do not experience as much absolute dollar cost inflation because they do not employ as many inputs.

APPENDIX T-6.A

THE MASSACHUSETTS RATE SETTING SYSTEM

THE MASSACHUSETTS RATE SETTING SYSTEM

The following appendix presents a detailed summary of the Massachusetts nursing home rate setting system for 1973. Changes which occurred in 1974 and 1975 are identified:

Reporting Requirements. Under the system, providers filed an annual report (form RSC-1) for the preceding calendar year on or before March 15th of each year. Thus, information for calendar year 1972 must have been submitted by March 15, 1973.

Basis for the Interim Rate. Historical cost data (for 1973 rates, 1971 data were used) were used to set the interim rates for each year. Year end adjustments using actual costs determined final payment rates. The following cost components (each on a per diem basis) were summed in arriving at a facility's interim rate:

1) ireasonable operating costs[1]
2) nursing costs (RNs and LPNs only)
3) administration and policy planning allowances
4) additional services and supplies, i.e. indirect restorative therapy not billed for directly
5) return on average equity capital
6) interest expense
7) depreciation
8) pension and retirement funds

The costs are taken from the RSC-1 cost reports. All costs in the first two categories (reasonable operating costs and nursing costs) are covered except those which are in excess of one standard deviation of the mean costs

[1] Reasonable operating costs included only non RN and non LPN nursing expenses and some "overhead" expenses: accounting and auditing expenses; approved education expenses for employees; advertising expenses (newspaper and other public media only); fringe benefits--group life, accident and disability insurance, with limits, and health insurance; rental expenses; leasehold interest on real estate; the value of services of non-paid workers; compensation for no more than two administrators-in-training; certain nursing home supplies not directly billed to any government unit; and various miscellaneous expenses.

of a representative sample of providers in the same group. For this purpose providers are grouped on the basis of certification type (4 categories) and bedsize (3 categories).

The amount of administrative and policy planning allowance is determined by a fixed schedule according to the levels of care provided by the facility type and bedsize categories. For the other components of costs, the Commission established maximum cost guidelines.

In addition to the above cost elements, a fixed percentage of net operating costs (excluding interest and depreciation) was added in calculating the interim rate. This cost adjustment factor was 15% in 1973.

As noted, each of the above (8) elements was computed on a per diem basis. The respective allowable costs or expenses were divided by the larger of (1) the facility's actual patient day census for 1971 or (2) 90% of the facility's licensed bed capacity, except that the administration and policy planning allowances and the return on average capacity were divided by the larger of (1) the provider's actual patient day census for 1971 or (2) 90% of the provider's total bed capacity.

Basis for the Final Rate of Payment. The determination of the final rate of payment involved much of the same type of calculations as used in computing the interim rates. However, final rates were based on actual costs as determined by a field audit of the books and records of the provider. The same (8) categories were examined. Occupancy penalties were still used and the same ceilings on reasonable operating costs and nursing costs still applied. However, in calculating the final rate, if a portion of a provider's nursing costs were disallowed and the provider had no disallowance for the reasonable operating costs category, then the disallowed nursing costs could be included up to an equal amount to the difference between the ceiling established for reasonable operating costs for the group into which the provider was classified.

The Appeals Process. Under certain conditions, a provider was allowed to file a petition to the Rate-Setting Commission requesting an adjustment in either the interim rate or final rate of payment. Valid reasons for the filing of a petition included (1) if there was an error in the calculation of the rate; (2) if the petitioner incurred or planned to incur a substantial capital expenditure; (3) if the petitioner had made or

planned to have made a substantial change in services or an expenditure in order to have most of the Department of Public Health (DPH) requirements; (4) if certain conditions were present in the transfer of ownership; or (5) if unforeseen increases in reasonable operating costs threatened the financial stability of the providers.

Changes to the System in 1974. There were no significant changes in the reimbursement regulations in 1974.

Changes to the System in 1975. There were a few noteworthy changes to the system in 1975. They are:

(1) Nursing costs were newly defined as compensation for nursing personnel including registered nurses, licensed practical nurses, rehabilitation, nursing assistants and nurses' aides.

(2) With regard to depreciation of buildings and equipment, an addition to the regulations did not permit reimbursement for building improvements in order to meet environmental and sanitation standards for certification, including the life safety code, or any improvement in excess of minimum compliance.

(3) The rate of return on average equity capital was increased from 11% to 11.9%.

(4) The bed capacity lower limit was raised from 90% to 93%.

(5) An allowance for costs incurred was made for initiating innovative and special programs.

APPENDIX T-6.B

MEANS AND STANDARD DEVIATIONS OF MASSACHUSETTS DATA

Table T-6.B.1: MEANS AND STANDARD DEVIATIONS FOR 1973 (based on 453 cases)

Variable	Mean	Standard Dev.
Capacity		
BEDS	66.5166	42.6410
Utilization		
OCC	94.3896	9.8791
Patient Mix		
PLE59	0.0662	0.0887
P8089	0.4152	0.1470
POLD	0.1292	0.0832
Sex		
PFEM	0.7321	0.1852
Admission Source		
PADNH	0.1268	0.1336
PADHO	0.2178	0.1743
Conditions		
BEDRIDDEN	0.0347	0.0734
DEAF	0.0232	0.0459
BLIND	0.0281	0.0295
DISTURBED	0.1534	0.2334
RETARD	0.0438	0.0882
EPISEPSY1	0.0203	0.0323
PALSY	0.0040	0.0109
Services		
PT	0.9338	0.2490
OT	0.7108	0.4539
SPHT	0.7263	0.4464
COUNSEL	0.8256	0.3799
Case Flow		
ADMISPD	0.0024	0.0033
Location		
REG1	0.1148	0.3191
REG2	0.1369	0.3441
REG3	0.0728	0.2602
REG4	0.1214	0.3270
REG5	0.1369	0.3441
REG7	0.1634	0.3701
REG8	0.0993	0.2994

Table T-6.B.1: MEANS AND STANDARD DEVIATIONS FOR 1973 (based on 453 cases) (cont'd)

Variable	Mean	Standard Dev.
Regulation Payment Proportion		
PCPRIV	0.2804	0.2295
Licenses and Certificates		
CERCARE	0.1258	0.3320
PBSNF	0.2493	0.3429
PBRES	0.0135	0.0792
Facility Characteristics		
NONPROF	0.0618	0.2411
MIXED	0.3091	0.4626

Table T-6.B.2: MEANS AND STANDARD DEVIATIONS FOR 1974 (based on 468 cases)

Variable	Mean	Standard Dev.
Capacity		
BEDS	70.363	46.191
Utilization		
OCC	95.328	46.191
Patient Mix		
PLE49	0.022	0.042
P5064	0.084	0.088
P6569	0.079	0.070
P8089	0.395	0.139
POLD	0.136	0.086
Sex		
PFEM	0.739	0.184
Conditions		
WALK	0.276	0.199
DEAF	0.020	0.035
BLIND	0.030	0.030
PSYCH	0.106	0.178
RETARD	0.045	0.086
EPILEPSY2	0.022	0.030
EPILEPSY3	0.002	0.009
PALSY	0.004	0.013
CANCER	0.039	0.043
DIABETES1	0.128	0.070
Services		
PT	0.573	0.495
PNVRC	0.879	0.229
PONEB	0.075	0.138
PTHREEB	0.313	0.282
PFOURB	0.263	0.261
Case Flow		
ADMISPD	0.002	0.002
Location		
REG1	0.115	0.320
REG2	0.120	0.325
REG3	0.073	0.260
REG4	0.139	0.346
REG5	0.132	0.339
REG7	0.156	0.363
REG8	0.088	0.283

Table T-6.B.2: MEANS AND STANDARD DEVIATIONS FOR 1974 (based on 468 cases) (cont'd)

Variable	Mean	Standard Dev.
Regulation		
Payment Proportion		
PCPRIV	0.239	0.218
Licenses and Certification		
PBCARF	0.076	0.211
PBDF	0.173	0.280
PBRES	0.021	0.109
NONPROF	0.077	0.267
CHAIN	0.083	0.277
MIXED	0.348	0.477

Table T-6.B.3: MEANS AND STANDARD DEVIATIONS FOR 1975 (based on 438 cases)

Variable	Mean	Standard Dev.
Capacity		
BEDS	71.8265	46.8426
Utilization		
OCC	95.3447	4.6481
Patient Mix		
PLE49	0.0212	0.0463
P5064	0.0888	0.0959
P6569	0.0772	0.0677
P8089	0.4048	0.1515
POLD	0.1357	0.0821
Sex		
PEEM	0.7471	0.1743
Admissions Source		
PADNH	0.1196	0.1165
PADHO	0.1761	0.1518
Conditions		
WALK	0.2381	0.1939
DEAF	0.0204	0.0464
PSYCH	0.1009	0.1721
RETARD	0.0147	0.0667
EPILEPSY2	0.0286	0.0351
EPILEPSY3	0.0010	0.0074
PALSY	0.0046	0.0148
CANCER	0.0182	0.0257
DIABETES2	0.0663	0.0613
DIABETES3	0.0837	0.0613
Services		
PT	0.2237	0.4172
PNURC	0.8170	0.2678
PONEB	0.0728	0.1310
PTHREEB	0.3112	0.2780
PFOURB	0.2691	0.2519
Case Flow		
ADMISPD	0.0019	0.0024
Location		
HSA1	0.1096	0.3127
HSA2	0.1393	0.3466
HSA3	0.0708	0.2567
HSA5	0.1461	0.3536
HSA6	0.1279	0.3343

Table T-6.B.3: MEANS AND STANDARD DEVIATIONS FOR 1975 (based on 438 cases) (cont'd)

Variable	Mean	Standard Dev.
Area Descriptors		
WAGE	1.2361	0.0653
Regulation Payment Proportion		
PCPRIV	0.2300	0.2234
Licenses and Certificates		
PBCARE	0.0654	0.1949
PBDF	0.2007	0.3007
PBRES	0.0178	0.0962
Facility Characteristics		
NONPROF	0.0913	0.2884
NCHAIN	2.3402	5.2564
MIXED	0.3584	0.4801

TECHNICAL PAPER #7

Technical Paper #7

NURSING HOME COSTS IN INDIANA*

T-7.0 Introduction

This chapter presents the econometric analysis of the costs of nursing home care over a three-year period in Indiana. The approach to the cost function analysis applied to Indiana differs from the state analyses of New York and Massachusetts in that the unit of observation is taken to be the county rather than the nursing home itself. The remainder of the chapter is organized as follows. Section T-7.1 identifies the implications of this difference in the context of estimating a theoretical model applicable to individual facilities. Section T-7.2 discusses the type of regulatory system for nursing homes in effect over the three-year period examined: 1973, 1974, and 1975. Section T-7.3 identifies the source and nature of the data used in the Indiana study. Section T-7.4 discusses the types of variables actually used to estimate the theoretical cost model developed below. Section T-7.5 presents the empirical results, looking first at individual years, then integrating the results across the three years. Section T-7.6 concludes the chapter with a summary of the findings.

T-7.1 Theoretical Model

This section briefly discusses the hypothesized model of nursing home cost behavior in Indiana. The average operating cost model is similar to the one developed for the national analysis of nursing homes. Four basic types of variables are hypothesized as arguments in the function for average operating costs: measures of nursing home output, input prices, variables reflecting regulation, and variables which further characterize the production processes of facilities. In general terms the hypothesized model across the three years can be written as:

$$AOC = f(Q, P, R, F)$$

where

AOC = average per diem operating expenses

Q = vector of output related variables including measures of the scale of production, services offered, patient mix, and the turnover rate of patients

P = vector of input prices

R = vector of regulatory variables including certification status of the home, the payment source of patients, and changes in state regulations across the three years.

---

*by Gail Jensen

F = vector of additional facility characteristics describing production differences, including the types of ownership, whether or not the facility is part of a chain, and any changes in the bedsize of the facility.

A discussion of the motivation for including each of these types of variables will be be presented here. The reader is referred to the discussion of the model in Chapter 4 and Technical Paper T-3.

The analysis of the determinants of nursing home costs presented here is different from the National, Massachusetts or New York analyses in that the unit of observation is the county rather than the individual facility. Although the county is the unit of analysis we can expect the results to be representative of individual facilities, since in general, the estimated regression weights associated with a county level model will be unbiased estimates of the regression weights associated with a facility level model.[1] Thus, inference from the equation is similar to what it would be if in fact the home was the unit of observation.

It is noted that although the estimated regression coefficients are in general unbiased, the efficiency of estimated parameters decreases as intra-county variation among homes increases. In loose terms, this means that the ordinary least squares estimates will give a correct estimate, on average, for the parameters of the cost function applicable to individual facilities. In addition, the estimates will be closer to the true values of the parameters as the intra-county variation among homes decreases. Given the available data, there is no way to measure intra-county variation and so no way to test whether the variation among homes is small or large. However, the results reported here are suggestive of the actual cost functions for Indiana homes.

T-7.2 The Nursing Home Environment in Indiana

Indiana has had a prospective reimbursement system since 1970. The prospectively set rate for the time period examined here was based on annually submitted cost reports of the nursing homes which reflect costs as identified by the homes. Blue Cross, under a contract with the Department of Public Welfare, reviewed the nursing home cost reports and computed reimbursement rates.

---

[1] A technical justification of this result and those mentioned in the next paragraph is located in Kmenta (1971, p. 322-336).

More specifically, the system works as follows. First, for each home a per diem rate with some allowable retrospective adjustments (primarily in depreciation) was identified based on the facilities' reported costs. Then the facilities' rate was compared to the average rate for facilities of the same certification type (skilled or intermediate) and in the same designated market area.[1] If it was felt that a home's rate was too high, then it was adjusted downward to the maximum limit which was set as 20 percent higher than the group's average (the group being determined by certification and market area location). Facility rates could not be increased more than twice in any 12 month period. Our conversations with the Department of Pulbic Welfare suggest that everyone liked the system and that there is a harmonious relationship between the state, Blue Cross, and the nursing homes. However, the extent to which this harmonious relationship works to achieve the goals of public policy is unclear.

Coupled with the system for setting reimbursement rates, Indiana had a survey and certification program for ICFs and SNFs. Visits to nursing homes by registered nurses or specialists (including dieticians, pharmacists, physical therapists, and sanitarians) were made in an effort to maintain set certification standards pertaining to the quality of care provided to patients. All SNF visits were unannounced. There were seven full-time nurse surveyors (the number has been constant since 1974). Also, surveyors with a particular specialty were hired as consultants. For ICF facilities, surveyors made at least three visits annually. For a randomly selected 10 percent of the facilities, all three visits were unannounced, but for the remaining 90 percent, only two of the three visits were unannounced. In addition, unannounced visits were sometimes made to check on any complaint against a facility.

The Indiana nursing home regulatory system is described in detail for 1973, 1974, and 1975 in Appendix T-7.A. In regard to changes in regulation over the time period being examined, there were no major changes in the system from 1973 through 1974. However, there was a slight change in the system in 1975, namely that any new addition of facilities or equipment could only be depreciated on a straight line basis.

---

[1] For purposes of the Medicaid program the market areas are defined as the three Indiana Health Service Areas, which cut the state into three rectangular regions, northern, central and southern.

T-7.3 The Data

Three years of county level nursing home data (for 1973, 1974, and 1975) were provided by the Department of Public Welfare. The data base contained information on the number of nursing homes in the county, the average breakdown of costs by accounting components, the average occupancy rate, and the Medicaid occupancy rate, all by certification types of facilities. The described data were county averages for each group of facilities with a similar certification (ICF only, mixed ICF-SNF, or SNF only). The number of beds in each certification group, the average number of nursing hours per patient day, and the average wages of nursing home RNs, and of all nursing home workers were also reported for each county. The data was constructed from annual cost reports maintained by the Indiana State Department of Public Welfare.

The sample sizes for the three years differ slightly since in each year there were some counties with partially missing data.[1] For purposes of analysis, average per diem operating expenses were defined as total per diem expenses minus per diem mortgage and interest, per diem depreciation of building, equipment and fixtures, per diem lease/rental of building and equipment, per diem insurance expenses, and per diem property, state gross income, and federal gross income taxes.

All reported results are in "real" 1973 dollars.[2]

T-7.4 The Variables

The theoretical model of nursing home cost behavior indicated the types of variables which should enter the cost function. Unfortunately, not all of these variables were available for Indiana for the three time periods considered here; among those missing were measures of the quality of care, the mental and physical debility mix in the county, measures of patient turnover, a measure of the percentage of Medicare patient days, and data on ownership and other facility characteristics. Thus the ideal variables were proxied by the available variables. The following paragraphs examine the actual variables used. Hypotheses associated with the various variables are also discussed.

---

[1]The sample sizes for the three years were 78 for 1973, 86 for 1974, and 82 for 1975.

[2]To compute the nominal quantities, the inflation factor for 1974 is (147.7/133.1) = 1.1097, and for 1975, (161.2/133.1) = 1.2111. The price indices used for deflation were the medial care components of the consumer price index (base year, 1967) for the midwestern region.

T-7.4.1 Output Measures

It is hypothesized that the size and occupancy rate of a facility affects per diem operating expenses such that there are both economies of scale and economies of utilization. It is expected that, ceteris paribus, as the scale of production increases, cost declines, and for a given capacity, higher occupancy rates imply lower costs. Average daily census (ADC), i.e., annual patient days in 365 day units, is used to represent the scale of production. Both ADC and the occupancy rate (OCC) were entered into the model in a way which allowed for non-linearities, using the "grafting" procedure outlined in Technical Paper #2.

Services offered are a third dimension of nursing home output. It is hypothesized that the more services a facility offers, the greater its average operating costs, ceteris paribus. The only measures of service intensity available to the Indiana analysis were the percentage of operating expenses paid for physical therapy (PCTPT) and the percentage paid for counseling and other social services (PCTCOUN). It is expected that an increase in either of these two variables increases per diem costs. Unfortunately, data on other types of offered services (e.g. rehabilitative therapy, occupational therapy, and speech and hearing therapy) as well as ancillary services such as nasal feeding, enemas, catheterizations, and full bed baths) were not available. No data on the patient debility mix in the home or on the rate of patient turnover were available.

T-7.4.2 Input Prices

It is hypothesized that per diem operating costs are positively related to input prices, i.e., homes with higher input prices (e.g. homes located in urban areas) are expected to have higher per diem costs. Included in the model was the average county nursing wage among nursing home workers (NURSEWG) and a dummy variable (SMSA) indicating whether or not the county was part of an SMSA. It is expected that urban homes have higher costs.

T-7.4.3 Regulations

All analyses in this study used certification level to measure product differentials and to allow for the possibility that certification level (SNF/ICF) insures that a common minimum set of services is available, though homes may provide services in excess of the minimum. The certification variables are considered here as regulatory variables since SNF or mixed SNF-ICF beds will have to follow more stringent regulations than ICF

homes. The percentage of county beds which are SNF certified (PCTSNF) and the percentage of county beds which are both SNF and ICF certified (PCTMIX) were used to represent the certification mix of the homes in the county. In addition, partial data on the mix of patients in the county by payment source was available. The percentage of patients whose primary payment source was Medicaid (PCTCAID) was used to account for patient mix by payment source.

T-7.4.4 Additional Facility Characteristics

Certain facility characteristics such as ownership type, whether or not the facility is part of a chain, and whether or not the facility has changed recently its bed capacity may affect operating costs. Regretably, no data of this sort was available for the study.

Table T-7.1 contains the list of variables used for analysis and their definitions.

Appendix T-7.B contains the means and standard errors of the variables.

T-7.5 Results

The issue of a suitable functional form and the appropriate estimation technique are discussed first. Then the results for each of the three years are considered individually. For each year, the differential impacts of the types of causal variables identified are discussed. Next, the three equations are compared to each other in terms of similarities and differences.

T-7.5.1 Functional Form and Estimation Technique

There is a broad range of possible parametric cost function forms which allow for testing of the basic hypotheses set forth above (e.g., the presence of economies of scale and utilization, the positive relationship between input prices and costs, between urban status and costs, etc.). Due to the potential expense and intractability of reporting a very large set of results, the Indiana study followed the reasonable and general empirical strategy of fitting an ordinary least squares regression of costs on grafted segments[1] in both average daily census and the occupancy rate (so that slopes vary for different ranges of these two variables), with all other

---

[1] The methodological discussion in Technical Paper #2 presents a technical derivation of the grafting approach. A heuristic discussion of the approach is contained in Chapter 4.

TABLE T-7.1

VARIABLE DEFINITIONS: INDIANA ANALYSIS

AOC	weighted average across SNF, mixed, and ICF homes of operating costs per patient day;*
ADC	weighted average across SNF, mixed, and ICF homes of the mean daily census*
OCC	weighted average of the percentage of beds occupied*
PCTSNF	percentage of county beds which are certified as SNF beds
PCTMIX	percentage of county beds which are certified as both SNF and ICF
SMSA	1 = part of a standard metropolitan statistical area 0 = otherwise
PCTPT	weighted average of the percentage of operating expenses which paid for physical therapy*
PCTCOUN	weighted average of the percentage of operating expenses which paid for counseling and other social services*
AVNURWG	average facility nursing wage in the county (includes RNs, LPNs, and aids)
PCTCAID	weighted average of the percentage of Medicaid patients*
YEAR74	1 = data applicable for 1974 0 = N/A
YEAR75	1 = data applicable for 1975 0 = N/A

*weights defined as the proportion of county beds which are SNF certified, the proportion which are both SNF and ICF certified, and the proportion which are ICF certified.

variables entered in a linear fashion. This approach of grafting allowed the data to approximate one of many possible non-linear forms with regard to ADC and the occupancy rate. Since there was no a priori hypothesis regarding the functional form of operating expenses with respect to service intensity, input prices, and the regulatory variables, they were entered additively and linearly; the national analyses of multiplicative forms found that such functional forms neither contributed to the explanatory power of the regression nor effected the substance of the results.

T-7.5.2 Results for 1973

Turning to the cost equation for 1973, which is reported in Table T-7.2 the R^2 for this equation was 0.721. Seventy-eight counties were used for analysis. The main output variable was average daily census (ADC). Costs were modeled as a function of six grafted linear segments in this variable, ADC, the join points being 50, 100, 150, 200, and 250. With regard to this variable there was, as found in the national results, no definite pattern to the signs of the slopes across intervals, and also no definite pattern to the significance of the coefficients. Three consecutive intervals showed significant coefficients but the signs alternate: negative-positive-negative. (The actual magnitudes of the coefficients are also extremely small.) The occupancy rate was also fit in a grafted fashion; the join points being 80, 85, 90, 93, and 97. None of the slope coefficients for this variable were significant, and as with ADC, there was no definite pattern to the signs of the slopes. These results, that there are not significant economies of scale or utilization, support the national and other state findings.

The two service intensity variables used (the percentage of expenses allocated to physical therapy and the percentage allocated to social services) did not appear to be cost determinants since both variables were statistically insignificant. This result is not surprising since the two variables are such poor proxies for service intensity; rather than the percentage of operating expenses going to the corresponding service, it would have been desirable to use either the percentage of patients having received the service or the average number of times per patient the service was received within a time interval.

Table T-7.2

DETERMINANTS OF AVERAGE OPERATING COSTS FOR INDIANA COUNTIES; 1973 RESULTS

Explanatory Variable	Estimated Coefficient	Standard Error
Output: Scale		
CONSTANT	4.397	
AD1	0.058	0.040
AD2	-0.030xx	0.014
AD3	0.044xx	0.019
AD4	-0.071xx	0.030
AD5	0.027	0.026
AD6	0.0005	0.001
OC1	0.1002	0.023
OC2	-0.008	0.259
OC3	0.199	0.297
OC4	-0.032	0.353
OC5	-0.104	0.193
OC6	0.002	0.003
Output: Services		
PCTPT	-0.112	0.199
PCTCOUN	-0.560	0.450
Input: Prices		
NURSEWG	3.592xxx	0.941
SMSA	0.004	0.468
REGULATORY VARIABLES		
PCTSNF	0.071	0.056
PCTMIX	0.037xxx	0.007
PCTCAID	0.007	0.012
R^2	0.721	
Adjusted R^2	0.630	
F-ratio	7.925	
N	78	
Standard Error	1.475	

x significant at .10 level
xx significant at .05 level
xxx significant at .01 level

A strong positive statistical relationship was found between input prices as measured by the average county nursing wage (AVNURWG) and per diem operating expenses. It appears that counties which were part of an SMSA had about the same costs as the more rural counties as indicated by the positive sign but statistically insignificant coefficient of this variable.

Next, with regard to the 1973 regulatory variables, it was found that mixed SNF-ICF homes were significantly more costly than pure ICFs, and pure SNFs were slightly more costly than mixed ICF-SNFs. A unit increase in the percentage of mixed SNF-ICF beds in the county increased average county costs by approximately $0.037. A similar (one percentage point) increase in the percentage of pure SNF county beds increased costs by approximately $0.071, ceteris paribus. Finally, it appeared that Medicaid patients were only marginally more expensive than non-Medicaid patients as indicated by the positive but very insignificant coefficient of PCTCAID. (The associated test statistic is given by 0.325 F(1,58).)

T-7.5.3 Results for 1974

For the 1974 regression (see Table T-7.3) the results are somewhat different than those found for 1973, possibly due to the use of data on a different number of counties. Eighty-six counties were used for analysis. The equation for year 1974, reported in Table T-7.3 had an R^2 of 0.471. The only significant variable in the equation was PCTMIX, the percentage of county beds which were both SNF and ICF certified. With regard to the scale operations, ADC and the occupancy rate, there were no noteworthy differences from the 1973 equation, i.e., there was no evidence of either economies or diseconomies of scale and/or utilization. The two service intensity variables were again insignificant.

Surprisingly, input prices appeared as insignificant, although their positive relationship to costs was evident. SMSA classification was again insignificant but different in sign for 1974 (now negative).

Finally, with regard to the three regulatory varibles, mixed SNF-ICF facilities were significantly more expensive than ICF facilities (the omitted category). The coefficient of PCTSNF, the percentage of SNF beds in the county, was negative but insignificant. For 1974, non-Medicaid patient days were slightly more costly than Medicaid patient days as indicated by the negative but insignificant coefficient of PCTCAID.

Table T-7.3

DETERMINANTS OF AVERAGE OPERATING COSTS FOR INDIANA COUNTIES; 1974 RESULTS

Explanatory Variable	Estimated Coefficient	Standard Error
Output: Scale		
CONSTANT	10.495	
AD1	0.042	0.056
AD2	0.0009	0.017
AD3	-0.009	0.025
AD4	0.018	0.028
AD5	-0.0008	0.021
AD6	0.00008	0.001
OC1	0.014	0.030
OC2	-0.052	0.384
OC3	-0.044	0.541
OC4	-0.411	0.543
OC5	0.379	0.306
OC6	-0.124	0.520
Output: Services		
PCTPT	0.214	0.181
PCTCOUN	-0.204	0.344
Input: Prices		
NURSEWG	0.448	0.816
SMSA	-0.554	0.507
REGULATORY VARIABLES		
PCTSNF	-0.025	0.031
PCTMIX	0.031xxx	0.007
PCTCAID	0.017	0.016
R^2	0.471	
Adjusted R^2	0.318	
F-ratio	3.094	
N	86	
Standard Error	1.793	

x significant at .10 level
xx significant at .05 level
xxx significant at .01 level

T-7.5.4 Results for 1975

The regression results for 1975 (see Table T-7.4) were strikingly similar to those discussed for the 1974 equation. This was expected since the only change in nursing home regulation over the 1974-1975 period was the addition of a new requirement regarding the use of straight line depreciation in place of accelerated depreciation on new equipment or additional facilities. Eighty-two counties were used for analysis. The equation, reported in Table T-7.4 accounts for 54.4 percent of the variation in operating expenses. There were few noteworthy differences from the 1974 regression. Again, the only significant variables were the percentage of county beds which are both SNF and ICF certified. The only change in the sign and/or significance of the explanatory variables was in regard to the percentage of Medicaid patient days which had a positive sign in 1975 but again insignificant coefficient. In general, the coefficients in the 1975 equation are similar in absolute value to the 1974 coefficients.

T-7.6 Comparisons Across 1973, 1974, and 1975

The poor quality of data available for Indiana severely restricted the analyses; the results that were found can be seen only as a complement to the other state analyses, for they cannot stand on their own. To highlight the similarities and differences of the three equations across the years, 1973, 1974, and 1975, there are neither economies nor diseconomies of scale and/or utilization for any of the three years examined as indicated by the overall insignificance of the associated coefficients across the three years. The two service intensity variables (the percentage of operating expenses allocated to physical therapy and the percentage allocated to social services) were both insignificant in all years.

The coefficient of input prices had a positive sign but was insignificant in both 1974 and 1975. SMSA status was insignificant in each year and negatively related to costs in years, 1974 and 1975.

With regard to certification, a consistent finding across years was that the percentage of beds which were both SNF and ICF certified was positively and significantly related to costs indicating that mixed SNF-ICF homes were more costly than pure ICF homes. The percentage of beds which were SNF certified was insignificant (relative to the omitted categoy, the percentage of ICF beds) throughout the three years. The coefficient of this variable

Table T-7.4

DETERMINANTS OF AVERAGE OPERATING COSTS FOR INDIANA COUNTIES; 1975 RESULTS

Explanatory Variable	Estimated Coefficient	Standard Error
Output: Scale		
CONSTANT	12.2113	
AD1	0.0077	0 0523
AD2	0.0027	0.0176
AD3	-0.0145	0.0171
AD4	0.0136	0.0207
AD5	0.0070	0.0229
AD6	-0.0003	0.0013
OC1	-0.0166	0.0309
OC2	-0.4333	0.4146
OC3	-0.4557	0.5515
OC4	-0.0574	0.4913
OC5	-0.0272	0.2616
OC6	-0.2308	0.4680
Output: Services		
PCTPT	0.1288	0.1420
PCTCOUN	-0.2223	0.3119
Input: Prices		
NURSEWG	0.8608	0.6616
SMSA	-0.1594	0.4813
REGULATORY VARIABLES		
PCTNF	-0.0216	0.0285
PCTMIX	0.0278xxx	0.0080
PCTCAID	0.0021	0.0179

R^2	0.5448
Adjusted R^2	0.4053
F-ratio	3.9064
N	82
Standard Error	1.63037

x significant at .10 level
xx significant at .05 level
xxx significant at .01 level

which was positive, as expected, for 1973, was negative for both 1974 and 1975. There was no change in regulation from 1973 to 1974 which could account for the change in the sign of this coefficient.

Finally, costs were found to vary positively with the percentage of Medicaid patient days for the years 1973 and 1975, but vary negatively for the year 1974. The coefficient of this variable was insignificant in each year.

Because of the few changes in regulation across the time period examined, the hypothesis that the three annual "real" per diem operating cost functions were essentially the same, was tested. A three year merged sample regression was estimated (see Table T-7.5) where the explanatory variables were those used in each year's regression with the exception that two dummy year variables (for 1974 and 1975) were also included. Table T-7.5 contains the regression results. An F-test was used to assess differences in the cost functions associated with each year. The full hypothesis was that the overall structure of operating expenses was the same in each of the years 1973, 1974 and 1975. The appropriate statistic was calculated as 0.89122 N F(38,198). It was insignificant at a 0.10 significance level (where the required level was F(38,197) = 1.30. This suggests that the cost function was the same for the three years, i.e., that a pooled equation is an adequate econometric specification.

To summarize the results of the pooled equation, the coefficients for the two year dummies (YEAR 74 and YEAR 75) were negative and significant at the 99 percent level, suggesting that real operating expenses decreased, ceteris paribus. PCTMIX (the percentage of county beds which are certified as both SNF and ICF) was positively related to costs and statistically significant at the 99 percent level indicating that the dually certified beds are more costly to maintain than ICF beds; PCTSNF, however, remained insignificant. The one other notably significant variable was AVNURWG, which proxied input prices. As expected its coefficient was positive. Variables other than those mentioned here were insignificant.

Table T-7.5

DETERMINANTS OF AVERAGE OPERATING COSTS FOR INDIANA COUNTIES; POOLED RESULTS (1973, 1974, and 1975)

Explanatory Variable	Estimated Coefficient	Standard Error
Output: Scale		
CONSTANT	9.9653	
AD1	0.0508	0 0263
AD2	-0.0085	0.0089
AD3	-0.0009	0.0108
AD4	0.0035	0.0134
AD5	0.0019	0.0124
AD6	-0.0003	0.0007
OC1	-0.0028	0.0149
OC2	-0.0047	0.1756
OC3	0.0539	0.2259
OC4	-0.1513	0.2400
OC5	-0.0414	0.1216
OC6	-0.0033	0.0034
Output: Services		
PCTPT	0.0366	0.0903
PCTCOUN	-0.1945	0.1862
Input: Prices		
NURSEWG	1.2694xxx	0.4127
SMSA	-0.3856	0.2615
REGULATORY VARIABLES		
PCTNF	-0.0115	0.0179
PCTMIX	0.0344xxx	0.0039
PCTCAID	0.0051	0.0083
YEAR 74	-0.8610xxx	0.2684
YEAR 75	-0.7960xxx	0.3052
R^2	0.5272	
Adjusted R^2	0.4829	
F-ratio	11.8973	
N	246	
Standard Error	1.6222	

x significant at .10 level
xx significant at .05 level
xxx significant at .01 level

APPENDIX T-7.A

INDIANA STATE NURSING HOME REIMBURSEMENT SYSTEM

INDIANA STATE NURSING HOME REIMBURSEMENT SYSTEM*

T-7.A.1 Overview of Nursing Home Regulation in Indiana [1]

Indiana has had a prospective reimbursement system since 1970, but nursing homes were not required to submit the annual cost reports until December 31, 1973. The cost reports reflect the costs as identified by the nursing home. Until 7/1/76, these costs were used to compute the rate. Now, some costs may be adjusted before rates are computed. The adjustments, primarily in depreciation, will not be on the data tapes.

The Department of Public Welfare contracts with Blue Cross to review nursing home costs and compute reimbursement rates. Everyone likes the system and there appears to be harmonious relations among the state, Blue Cross and the nursing homes. There is none of the divisiveness and rancor one senses in New York. The situation may change. Indiana has recently tightened and codified their reimburesement policy.

The Department of Public Welfare does not maintain a list of nursing home rates. An on-line system contains the data by provider number but the data is not available.

There is a survey and certification program for ICFs and SNFs. All SNF visits are unannounced. Visits are made by surveyors who are registered nurses and by consultants on an as needed basis. Consultants are specialists such as dietitians, pharmacists, physical therapists, sanitarians, etc. There are seven nurse surveyors. This number has been stable since 1974.

The number of ICF surveyors has gradually grown in the past few years from nine to the present 13-1/3 (the 5 FTE survey children's homes). A surveyor makes at least three visits to each facility. The annual survey is announced except for a 10 percent random sample. Prior to 1975, all annual surveys were announced. All other visits are unannounced. In addition to the three yearly visits, there can be unannounced visits to check on any complaint against a facility. A surveyor assumes that a third of the ICFs will receive at least one complaint a year.

---

*by Gail Jensen

[1] Taken from a memorandum to Howard Birnbaum and Christine Bishop from Michael Sumner, November 16, 1977.

T-7.A.2 Reimbursement System In Effect In Indiana In 1973-1976

1. Annual financial reports required. A nursing home's fiscal year can end at any time within a given year and its data are included in that year's data tape.

2. Nursing homes can request rate increases at any time. Most homes do not seek a rate increase. Of those that do, very few, if any, make more than two requests per year. The effective date of any rate increase is 90 days from the date the Medicaid Fiscal Agent (Blue Cross) receives the rate request.

3. Geographic area limitations were judgmental. Nursing homes within a county and surrounding counties were compared. If Blue Cross felt that a nursing home's prospective rate was greater than it should be, the rate was reduced. Between 10 and 17 percent of the homes were given reduced rates. Even though no guidelines were established for any reduction, nursing homes went along with the decision.

4. Ancillary services and personal care items are excluded from prospective rates. Ancillary services are physical therapy, speech and hearing therapy, occupational therapy, inhalation therapy, pharmacy, billable medical supplies, x-ray, laboratory, and barber/beauty shop. Personal items include dental and physical hygiene products.

5. Interest expense is a recognized cost and could not be greater than the historical cost of the building and equipment.

6. Depreciation was allowable as long as a generally accepted method was used. Providers with facilities wholly or partially depreciated and mortgages nearly or completely paid were <u>not</u> allowed to impute a fair market value in lieu of depreciation and interest.

7. Charity and courtesy allowances are deductions from revenue and are not included as allowable costs.

8. Unusual or non-recurring items such as tax penalties, legal demages or suits, etc. are not allowable costs. Start-up costs are allowable if amortized rateably over a 60-month period beginning with the month in which the first patient is admitted.

9. Management compensation related to bed size.

10. Cost applicable to services, facilities, and supplies furnished to the nursing home by organizations related to be common ownership or control are includable in allowable cost at the cost to the related organization. Common ownership was defined as any ownership no matter what the percentage. About 20 percent of the nursing homes conduct business with related organizations.

11. Minimum nursing staffing requirements are 2.5 hours per patient day in SNFs and 1.5 hours per patient day in ICFs. Maximum reimbursement payment is based on 2.5 nursing hours worked (excludes sick, holiday, vacation) per patient day in SNFs and 2.0 hours worked in ICFs.

12. Profit was designed to generate a 7 percent profit after tax. This amounted to a maximum of 12.42 percent of allowable cost.

Changes To This System In 1975

1. Depreciation - any new addition of facilities or equipment can only be depreciated on a straight line basis.

Changes To This System In 1976

1. Rates will not be increased more than twice in any 12 month period. Rates are limited to fees or charges that are consistent with the market area limitation and which can be no higher than the charges made to the general public for such services. The maximum limits for routine services is 120 percent of the weighted averages for comparable facilities and services in the market area.

   The designation "market area" is broader than the county and surrounding counties which defined a nursing home's geographic area. The state is now divided along HSA lines. There are three areas: Northern, Central and Southern. Nursing homes are categorized by type (SNF, ICF) and market area.

2. Only straight line depreciation allowed. If a nursing home's annual cost report uses accelerated depreciation, the report is not changed, but Blue Cross adjusts the figures for rate setting. The maximum computation of depreciation for facilities and equipment is $12,050 per bed adjusted for the amount of change in the construction element of the Consumers Price Index, on a semi-annual basis.

3. Interest is an allowable cost if it is incurred on a loan made to satisfy a financial need of the provider and for a purpose related to patient care. The maximum rate of interest for permanent financing at the time of closing cannot exceed a rate of 1-1/2 percent above Federal Home Loan Mortgage Corporation Whole Loan Purchase Multi-Family Rate. In addition, interest cannot be at a rate in excess of what a prudent buyer would pay in an arms-length transaction. Interest expense on permanent financing will be an allowable cost on amounts up to 75 percent of an idependent appraisal of the facility at the time of construction or purchase.

4. Recognition of the costs related to total staffing requirements (nursing, dietary, housekeeping, administrative, etc.) is limited to 4.5 paid hours per patient day on SNFs and 3.5 paid hours per patient day in ICFs. Paid hours include vacation, sick, and holiday pay.

5. In business conducted with related organizations, common ownership has been defined to be at or above 10 percent of the ownership.

6. As a management incentive, when a provider's occupancy exceeds 95 percent, the provider's rate will be computed utilizing the 95 percent occupancy level for fixed ownership costs (depreciation, interest, property tax and insurance, and maintenance).

7. The maximum allowable increase in a provider's average Medicaid rate will not be greater than nine percent in any 12 month period. The profit will be included in the approved charge as a fixed fee added to the allowable costs. This fee will be reviewed and updated by the Indiana State Department of Public Welfare. The after tax profit goal is seven percent.

APPENDIX T-7.B

MEANS AND STANDARD ERRORS OF INDIANA DATA

MEANS AND STANDARD ERRORS OF INDIANA DATA

Table T-7.B.1: MEANS AND STANDARD ERRORS FOR 1973 (based on 78 cases)

Variable	Mean	Standard Dev.
Output: Scale		
AOC	15.512	2.428
ADC	140.012	185.615
OCC	91.181	21.237
PCTSNF	0.402	3.341
PCTCAID	55.275	15.304
PCTMIX	41.342	40.837
PCTPT	2.249	0.972
AVNURWG	2.045	0.256
SMSA	0.282	0.452
PCTCOUN	0.915	0.453

Table T-7.B.2: MEANS AND STANDARD ERRORS FOR 1974 (based on 86 cases)

Variable	Mean	Standard Dev.
Output: Scale		
AOC	14.718	2.173
ADC	171.757	242.095
OCC	69.676	22.373
PCTSNF	1.883	7.520
PCTCAID	55.621	13.728
PCTMIX	40.073	36.144
PCTPT	2.660	1.276
AVNURWG	2.160	0.300
SMSA	0.255	0.438
PCTCOUN	0.077	0.649

Table T-7.B.3: MEANS AND STANDARD ERRORS FOR 1975 (based on 82 cases)

Variable	Mean	Standard Dev.
Output: Scale		
AOC	14.842	2.114
ADC	169.416	217.860
OCC	72.983	23.065
PCTSNF	1.701	7.732
PCTCAID	57.66	12.498
PCTMIX	36.793	36.264
PCTPT	2.715	1.419
AVNURWG	2.340	0.330
SMSA	0.243	0.432
PCTCOUN	1.183	0.649

Table T-7.B.4: MEANS AND STANDARD ERRORS FOR POOLED YEARS (based on 246 cases)

Variable	Mean	Standard Dev.
Output: Scale		
AOC	15.011	2.256
ADC	160.827	13.796
OCC	74.454	39.502
PCTSNF	1.353	6.582
PCTCAID	56.192	13.847
PCTMIX	39.382	37.627
PCTPT	2.548	1.253
AVNURWG	2.184	0.320
SMSA	0.260	0.439
PCTCOUN	1.061	0.601
YEAR 74	0.349	0.477
YEAR 75	0.333	0.472

TECHNICAL PAPER #8

TECHNICAL PAPER #8

A MULTI-EQUATION MODEL OF NURSING HOME COST BEHAVIOR*

T-8.0 Introduction

In this paper the 1972-1973 NCHS national data are used to explore a more complex model of nursing home behavior and examine whether or not it has any different implications for the design and evaluation of nursing home reimbursement alternatives. If cost were the only criterion of interest, it might be sufficient to estimate single-equation models of nursing home cost behavior--as done in Chapter 3. However, cost *per se* is not the only aspect of nursing home behavior about which the public is concerned. For example, the nursing home output may be regarded as being multi-dimensional in nature and composed of (1) the number of patient-days of care and (2) the "quality" or service intensity associated with each day. Whereas the cost functions already estimated describe the effect of reimbursement and regulation on the operating cost per patient-day, any impact on service intensity cannot be seen from these relationships. Thus, the reimbursement scheme that might emerge from a cost modeling effort alone could entail unknown and unacceptable reductions in service intensity. Moreover, it could also entail unintended and undesired consequences along other dimensions of interest (e.g., occupancy and access).

Furthermore, failure to take account of these effects in single-equation cost estimation can lead to serious misinterpretation of the results. For example, if service intensity were included in the cost equation--as it has been in single-equation estimation--and service intensity were found to be related to reimbursement, then the coefficients elicited for reimbursement variables in the cost equation would not give their *total* effect on cost. They would give only their direct effects. The indirect effects--i.e., the cost effects of reimbursement-induced service intensity differentials--would be reflected in the service intensity coefficients and thus would not be apparent from single-equation estimation.

This paper investigates such prospects--that is, it investigates whether or not nursing home operators react to reimbursement regulation along

---

*by A. James Lee, Howard Birnbaum, Christine Bishop, and Gail Jensen

relavant behavioral dimensions other than cost. In particular, it models empirically the effects of reimbursement and regulation on the following variables of interest:

(1) nursing hours per patient-day;

(2) an index of rehabilitative-type services;

(3) the occupancy rate;

(4) the mix of public-private patients; and

(5) the rate charged to private patients.

However, these variables are themselves inter-related and it is not possible to estimate reduced-form relationships for them without first specifying the underlying structural (or simultaneous) model.[1] Furthermore, such a structural model must itself be estimated and favorably evaluated if the reduced-form relationships that derive from it are to have validity.

Because the discussion here is by nature quite complex, a brief summary of results is useful. In brief, the results of the multi-equation model underline the importance of simultaneity issues in analyzing nursing home behavior. For example, it was found that an increase in service intensity increases operating costs, but it also increases occupancy rates and the percentage of private patients in a home--both of which lower costs, ceteris paribus. Furthermore, regulation and rate setting have a range of effects on variables other than cost. In particular, reimbursement variables also affect nursing and service intensity, the rates charged to private patients, and the public-private mix. For example, prospective reimbursement leads to an increase in the number of nursing hours per patient day and services provided to patients. It appears that retrospective and flat rate reimbursement systems which lead to a reduction in costs, also lead

---

[1] The structural form of a model is the set of behavioral relations--i.e., the ones derived from economic theory. The reduced form of a model is derived from the structural form and is the set of relations expressing the "equilibrium" impact of each exogenous variable on each of the variables being modeled, taking account of "feedback" or "multiplier" effects.

to a reduction in service intensity and access to public patients. Further, prospective reimbursement leads to an increase in the private rate and a decrease in the relative private-to-public mix of patients. Several other results should be noted. The demand for nursing home care by private-pay patients is not price inelastic; and further, private-pay patients are paying higher rates than public patients to begin, and cost controlling reimbursement features increase the differential. Further, private patients appear to subsidize public patients. There is no evidence that occupancy penalties affect the occupancy rate. Finally, there appears to be only a weak relationship between patient condition variables and service intensity and cost of care.[1] In sum, rate setting cannot proceed by simply analyzing the costs of nursing homes. Rather, a more comprehensive understanding of the nursing home environment is necessary if unintended and undesirable consequences are to be avoided.

The remainder of this paper is organized as follows. The structural model is developed in the next section, Section T-8.1. Structural estimates are reported in Section T-8.2. Reduced-form estimates are presented and discussed in Section T-8.3.[2] Finally, in Section T-8.4 the implications for design of a reimbursement scheme are discussed. Since, by its nature, the discussion is very technical, readers interested only in the conclusions may skip to the concluding section.

T-8.1 The Model

The multi-equation simultaneous system of equations developed herein describes nursing home market behavior in the short to intermediate run over which bed-size is fixed. The model is unusual inasmuch as it attempts to model supply and demand behavior with respect to individual nursing homes rather than within defined catchment or market areas. While a catchment area model would have been more conventional, the requisite data were not available.[3] Furthermore, this study is explicitly interested in

---

[1] This last result is subject to some justification because of an error-in-variables problem and the absence of measures of quality.

[2] If the reader is willing to take the underlying conceptual model on faith, Sections T-8.1 and T-8.2 can be skipped.

[3] Inasmuch as the NCHS data includes only a sample of nursing homes, it is not possible to aggregate information from all facilities in each catchment area. The sample is representative of the nation, but not of smaller geographic areas.

the predictability of individual nursing home behavior inasmuch as individual homes are the units for rate setting.[1]

The structural model encapsulates the key behavioral relationships identified from theoretic discussion in the literature review (Technical Paper T-2). Nevertheless, the specification of any analytic model, especially a multi-equation model, is highly judgmental--part art and part science. For this reason, a variety of alternative specifications might conceivably be equally compelling. Furthermore, any model must necessarily abstract from reality if it is to provide a useful and estimable description of it. Thus, many of the assumptions implicit in the specifications below are approximations of what is true, reflecting our judgement that certain relationships are more important than others given the data and other resource constraints of this study.

In any event, the model focuses on the six behavioral parameters noted above. Whereas the underlying hypothesis that regulation impacts these variables does not emerge herein from a formal model of nursing home behavior, nursing home operators and rate setters themselves suggest the importance of such effects. Furthermore ad hoc-type theoretic arguments lends support that regulatory variables impact on the behavior of nursing homes; it is the fundamental assumption underlying regulation that the regulations will induce changes in behavior. Consider now each of the behavioral variables to be examined. The nursing hours per patient-day and index of rehabilitative service variables--each measures a distinct service-quality dimension of nursing home care. It is hypothesized that cost-based reimbursement offers indirect incentive to increase the service intensity inasmuch as a greater "profit" accrues to more costly days.[2] Services are negatively-related to reimbursement control, since if costs are not met, the nursing home operator will have to curtail services provided or incur a loss. In addition, improvement of the quality or service intensity

---

[1] We are not aware that our facility-level specification suffers any serious econometric problems; the specification is merely unconventional for a multi-equation market model.

[2] This follows the typical instance wherein profit is a function of equity, and more services typically rely on increased capital expenditures.

of care may itself be an objective of some nursing home operators, especially in nonprofit facilities.

The occupancy rate is a straightforward measure of the quantity of nursing home care--i.e., quantity relative to capacity--provided by each nursing home. It is hypothesized that the quantity of patient care supplied by existing nursing homes varies directly with the overall generosity of public reimbursement--which is to say little more than that the supply curve is upward sloping.

The public-private patient mix, as reflected by the percentage of private patients variable, is an inverse measure for the relative access of public patients to nursing home care; the hypothesis is that the percentage of private patients increases with the ratio of private and public prices--or that, ceteris paribus, the lower the public rate of reimbursement, the larger is the private sector component since operators have relatively less of an inducement to take public patients. (Since the data indicate that a larger private sector component corresponds to a longer waiting list for admission, the public-private mix variable corresponds with reduced access for public patients.)

Finally, the per-diem rate charged by nursing homes to private patients may or may not be related to public reimbursement policies. The private and public markets for nursing home care are obviously not independent of one another, but the nature of interdependency has not been investigated empirically. For example, it is possible that the private rate subsidizes public patients or that a higher public rate also implies a higher rate for private patients. We shall test for such effects.

In specifying the relationships to be estimated in testing these informal hypotheses, one immediately confronts serious methodologic and econometric problems. First of all, given the general prevalence of cost-based reimbursement, it was not clear a priori that the rate of public reimbursement was itself a meaningful exogenous variable, inasmuch as the rate per se is not usually set by the public rate setter but rather emerges from the process prescribed for rate setting. That is, characteristics of the reimbursement system itself were thought to be the more relevant variables in terms of assessing the effects of alternative reimbursement alternatives. Furthermore, and perhaps more importantly, preliminary analysis

found that the public rate variable is a near-perfect linear combination of other variables in the model. Unfortunately, the implications of this finding are not altogether clear. Either the rates being set conform very closely to what is happening in the nursing home or, alternatively, the rates are having a monumental influence upon what happens in nursing homes. In any event, a public rate variable is necessarily excluded from the multi-equation model reported herein. (However, to the extent meaningful, we shall also comment on the implications of using a public rate variable.).

Although a variety of alternative models were also explored, and these will be mentioned as appropriate, the structural model actually reported herein can be broadly represented as follows:

The Average Operating Cost Function

$$(1) \quad AOC = f\ (Q,\ S_1,\ S_2,\ \frac{Q_p}{Q_p + Q_G},\ \frac{Q}{365 * BEDS},\ P_p,\ F,\ Y,\ P_{input},\ R,\ M)$$

The Two Service Intensity Equations

$$(2) \quad S_1 = f_2\ (Q,\ S_2,\ \frac{Q_p}{Q_p + Q_G},\ F,\ Y,\ P_{input},\ R,\ M)$$

$$(3) \quad S_2 = f_3\ (Q,\ S_1,\ \frac{Q_p}{Q_p + Q_G},\ F,\ Y,\ P_{input},\ R,\ M)$$

The Demand Equation (In Occupancy Form)

$$(4) \quad \frac{Q^D}{365 * BEDS} = f_4\ (S_1,\ S_2,\ P_p^D,\ \frac{Q_p}{Q_G + Q_p},\ F,\ R,\ M)$$

The Inverse Supply Function

$$(5) \quad P_p^S = f_5\ (Q,\ S_1,\ S_2,\ \frac{Q}{365 * BEDS},\ F,\ P_{input},\ R,\ M)$$

The Relative Supply Function

$$(6) \quad \frac{Q_p}{Q_G + Q_p} = f_6\ (Q,\ S_1,\ S_2,\ P_p,\ F,\ R,\ M)$$

Identities

$$(7) \quad p_p = P_p^S = P_p^D$$

$$(8)\quad \frac{Q_G}{Q_p + Q_G} = \frac{1-Q_p}{Q_p + Q_G}$$

$$(9)\quad Q^d = Q^S = Q$$

Where AOC is average operating cost per patient-day,

Q is the number of patient-days provided,

Q_G is the number of patient-days provided to government patients,

Q_p is the number of patient-days provided to private patients,

$\frac{Q_p}{(Q_p + Q_G)}$ is the fraction of total patient-days provided to private patients--equal to one minus the fraction of patient days provided to government patients,

S_1 and S_2 are indicators of service intensity or "quality",

BEDS is the number of beds in the nursing home,

$\frac{Q}{365 * BEDS}$ is the average occupancy rate,

P_p is the price per patient-day charged to private patients,

F is a vector of facility characteristics (e.g., profit/non-profit and ICF/SNF),

Y is a vector of patient characteristics,

P_{input} is a vector of input prices (e.g, wage levels),

R is a vector of characteristics describing reimbursement and other regulatory criteria (e.g., life-safety code requirements),

M is a vector of market characteristics (e.g., urban/rural, income level and nursing bed/population ratio),

and where superscript "d" denotes demand,

and superscript "s" denotes supply.

Analytically, the above set of equations can be described as a "block-recursive" system of equations. Equations (2) through (6) constitute a block of five equations that must be solved simultaneously because the dependent variables--S_1, S_2, $\frac{Q}{365 * BEDS}$, P_p and $\frac{Q_p}{Q_p + Q_G}$ are endogenously determined. That is, each of these dependent variables is either directly or indirectly a reciprocal function of the other dependent variables. For example, S_1 is a function of S_2 is a function of S_1. If identified, such a simultaneous-equations system can be solved using multi-equation estimation procedures. However, equation (1)--the average operating cost (AOC) function--is recursive and can be estimated using ordinary-least-squares since all independent variables are predetermined--i.e., they are either exogenous to the system or do not depend upon AOC in equations (2) to (6).

The model reflects the basic market structure implied by the assumption that the nursing home "output" is a composite of patient-days (Q) and service intensity (S), where P reflects prices. We have below in (i)-(iv) representations for the demand (1) and supply (5) of these two nursing home outputs, and in (v) and (vi) the equilibrium conditions that supply equals demand.

(4-10) (i) $Q^d = D, (P, S)$

(ii) $S^d = D_2 (P, Q)$

(iii) $Q^s = S_1 (P, S)$

(iv) $S^s = S_1 (P, Q)$

(v) $Q^s = Q^d = Q$

(vi) $S^s = S^d = S$

The Service Intensity Equations--(2) and (3)--are semi-reduced form relationships derived from (ii), (iv) and (vi) above. The Demand Equation (4) is meant to reflect the relative intensity of patient-days demand in a particular nursing home--i.e., a "small market" adjunct to (i) above. Equation (5) is simply the inverse of (iii) and equation (9) is the same as (v). Nevertheless, this market structure does not provide a complete guide to specification of the model. The behavioral situation is further complicated inasmuch as the patient-day output (Q) is divided between public and private

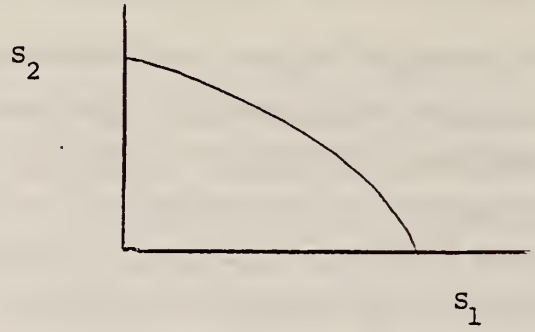
patients. Equation (6) is introduced for the purpose of modeling in a more ad hoc fashion the allocation of output between public and private patients.

Consider now the specifications for each of the system equations and the rationales for hypothesizing inter-relationship of dependent variables. Recall that the Average Operating Cost Function is not part of the simultaneous-equations set. Its conceptual representation and underlying hypothesis are much the same as those indicated in Chapter 3's single-equation cost estimation. However, potential differences are several. For example, the single-equation cost estimation did not include nursing hours per patient-day as one of the S-service intensity variables. This variable was excluded precisely because of a concern that it was actually endogenous, as is hypothesized here. A reduced-form cost function must therefore also be estimated excluding this and other endogenous variables. Thus, it is hypothesized that cost varies directly with service intensity. Similarly, as before, it is hypothesized that average cost decreases with capacity utilization or occupancy--also an endogenous variable in this setting.

Two Service Intensity Equations were indicated with services identified as S_1 and S_2. Operationally, these two service dimensions are (somewhat arbitrarily) defined as (1) nursing hours per patient-day and (2) an index of rehabilitative-type services. No doubt, the reader could specify other relevant aspects of service intensity and quality for inclusion in the model. However, preliminary investigation found that the multi-collinearity and identifications problems associated with a larger number of endogenous service variables confounded meaningful estimation of the multi-equation system. In each of the two service intensity equations, the one service variable is indicated as a function of the other--i.e., S_1 is a function of S_2 in equation (2) and S_2 is a function of S_1 in equation (3). The reasoning is as follows.

Ceteris paribus, especially given the overall level of reimbursement, a nursing home cannot provide more of one service without providing less of the other. In a sense, each nursing home ultimately confronts a

budget constraint and must operate along a production-possibility frontier (or curve) such as portrayed below:



That is, ceteris paribus, a trade-off is hypothesized to exist between S_1 and S_2 and hence the hypothesis is that each right-hand-side service variable coefficient will be negative in the other service equation. The percentage private patients variable $(\frac{Q_p}{Q_p + Q_G})$ is included in both service relationships to control for potential differences in public/private preference. The general presumption has been that private patients demand greater service intensity.

The occupancy equation (4) is represented as a Demand Equation. It is to reflect the level of catchment area demand for nursing home care within a particular facility.[1] The input price vector (P_{input}) is excluded, as it always is in specifying a demand-type relationship. The patient characteristics vector (Y) is also excluded inasmuch as this relationship is portraying community-level demand and the potential patient-mix (e.g., debility levels) are assumed not to vary much from community to community--or at least not as much as they do at the individual facility level. (Of course, this may not be true at the "margin" inasmuch as incremental demand then may depend upon the characteristics of patients already in the home. That is, a prospective male patient may not wish to enter a home that is predominantly female.) The service intensity variables (S_1 and

---

[1] Occupancy is merely the quantity "demanded" divided through by a multiple of bed-size to control for, in some sense, differences in the market size.

S_2) are included because they are, along with Q, dimensions of output. The average private rate (P_p) is the demand price variable for private patients. (Public demand is assumed to be price inelastic--i.e., insensitive to price--since no cost is incurred by Medicare or Medicaid patients.) Finally, the percentage-private variable ($\frac{Q_p}{Q_p + Q_G}$) is included to control, in the best available manner, for regional variation in the relative intensity of public and private demand. The general presumption is that private patients are willing to pay a premium for access and not have to wait for admission. Since the likelihood of having to wait varies directly with the occupancy level, an inverse relationship between occupancy and the percentage-private (endogenous) variable is hypothesized.

The private rate equation is meant to be an Inverse Supply Function. That is to say, it expresses the supply price as a function of the quantity-supplied. For this reason the vector of market variables (M) is excluded. As usual, it is hypothesized that the supply price varies directly with the quantity supplied--i.e., that it varies directly with occupancy. Similarly, it is hypothesized that a greater service intensity implies a higher price.

The Relative Supply Function has the percentage of private patients as its dependent variable and is thought to reflect the relative access of public and private patients to nursing home care. All other things equal, as the fraction of patient days provided to private patients increases, approaching one in the limit, increasing numbers of potential public patients are likely to end up on a waiting list. This assumes that private patients "bid up" the private price of care until their demand is met and that the adequacy of care provided to public patients depends upon the willingness of the public sector to bid against the private sector for available capacity. The usual presumption is that excess demand--as manifest by waiting lists--exists because the public sector is not willing to bid a sufficiently high price for its constituency. The patient characteristics vector (Y) is excluded from this relationship because patients are not inherently either public or private. The service intensity variables (S_1 and S_2) are included to control for differential public and private preferences. It is anticipated that private patients demand a greater service intensity. Ceteris paribus, it is hypothesized that a greater private price is required

to induce a nursing home to supply the private sector. That is, the nursing home is seen as being a "price maker" with respect to private patients.

The identities in equations (4) to (8) are straightforward and assume equilibrium in both the private and public components of the nursing home market. As already acknowledged, this latter assumption is clearly unrealistic. We sought at one point to portray a disequilibrium market and used a waiting list variable. However, we concluded that this could not be done in an operationable fashion without modeling in far greater detail the relationship between private and public demand components.

Turning now to the empirical specifications, the variables included in each relationship in the multi-equation system are indicated in Table T-8.1. The variables are defined in Table T-8.2; measured standard deviations for these variables are in Appendix 3-A. These specifications conform broadly to the conceptual guidelines indicated above. The exceptions stem from the necessity of excluding certain variables from relationships in order to achieve system identification. The rationales for omitting these variables as well as the hypotheses with respect to included variables, are discussed along with the results in order to avoid a lengthy and ultimately redundant discussion in this section.

The structural model is estimated and discussed in the next section, Section T-8.2. This is done primarily to validate the overall model--i.e., to demonstrate that the structural specification is an analytically meaningful one. However, it is more relevant for policy purposes to estimate and examine the reduced-form equations in Section T-8.3. All endogenous variables are omitted in estimating reduced-form relations, and the reduced-form coefficients give the equilibrium impact of a change in any exogenous variable, including the rate setting rules, on any endogenous variable.

T-8.2 The Structural Equations

The multi-equation structural model specified above has been estimated in Table T-8.3. The average operating cost (AOC) relationship was estimated by ordinary-least-squares. The other five relationships form a simultaneous-equations system, as discussed above, and were estimated by two-stage-least-squares techniques.[1] The dependent or endogenous variables in these relationships are as follows:

---

[1] The computer-generated R^2's are not valid for equations estimated by this procedure and have not been reported for this reason. However, the R^2 is valid for the OLS-estimated AOC (average operating cost) equation. It equals .558, compared to .585 from single-equation estimation.

Table T-8.1

Multi-equation Variable Specification

Independent Variable	Dependent Variable AOC	NHRPD	INDEX	OCCEN	PRIVRATE	PCTPRIV
NHRPD	X	--	X	X	X	X
INDEX	X	X	--	X	X	X
OCCEN	X	--	--	--	X	--
PRIVRATE	--	--	--	X	--	X
PCTPRIV	X	X	X	X	--	--
BEDS	X	X	X	X	X	X
CHANGE	X	--	--	X	--	--
ADMISPD	X	X	X	X	X	X
ICF	X	X	X	X	X	X
ICFSNF	X	X	X	X	X	X
PROFIT	X	X	X	X	X	X
CHAIN	X	X	X	X	X	X
PCONTRCT	X	X	X	--	X	X
PONEB	X	X	X	X	X	X
PTHREEB	X	X	X	X	X	X
AGE	X	X	X	--	X	--
PFEM	X	X	X	--	X	--
BEDPD	X	X	X	--	X	--
BEDD	X	X	X	--	X	--
PNOWDIG1	X	X	X	--	X	--
PNOWDIG2	X	X	X	--	X	--
PNOWDIG3	X	X	X	--	X	--
PNOWDIG4	X	X	X	--	X	--
PNOWDIG5	X	X	X	--	X	--
PNOWDIG6	X	X	X	--	X	--
PNOWDIG7	X	X	X	--	X	--
PNOWDIG8	X	X	X	--	X	--
PNOWDIG9	X	X	X	--	X	--
PNOWDIG10	X	X	X	--	X	--
PNOWDIG11	X	X	X	--	X	--
PNOWDIG12	X	X	X	--	X	--
PNOWDIG13	X	X	X	--	X	--
DEATH	X	X	X	--	X	--
PM1	X	X	X	--	X	X
PM2	X	X	X	--	X	X
LIM	X	X	X	--	X	X
OCCUPEN	X	X	X	X	X	X
UNIF	X	X	X	--	X	--
LSC	X	X	X	X	X	--
MA	X	--	--	--	--	X
CON	X	--	--	X	--	--
STAFFA	X	X	--	--	--	--
STAFFB	X	X	--	--	--	--
SMSA	X	X	X	X	X	X
WAGE	X	X	X	--	X	--
PCI	X	X	X	X	--	X
CONC	X	--	X	X	--	X
BEDPOP	X	X	X	X	--	X
(Constant)	X	X	X	X	X	X

Table T-8.2

VARIABLE DEFINITIONS: NATIONAL ANALYSIS

AOC	Operating costs per patient day, i.e., all annual non-capital costs per patient day that could be avoided ony by completely closing the nursing home divided by the number of annual patient days provided
NHRPD	Average number of nursing hours provided per patient day
INDEX	Index of services provided by nursing home, equal to average percentage of patients receiving one or more of the following services in the past seven days: counseling, nasal feeding, blood pressure reading, enema, catheterization, full bed bath, bowel and bladder retraining, application of dressing, temperature-pulse respiration, irrigation, intravenous feeding
OCCEN	Percentage of beds occupied by patients
PRIVRATE	Average per diem rate charged for private patient
PCTPRIX	Percentage of total patient days provided to private patients
BEDS	Number of beds in facility
CHANGE	1 = Bed size of home changed in past 12 months 0 = Otherwise
ADMISPD	Number of patients admitted in 1972 divided by annual patient days
ICF	1 = Type of certification is ICF only 0 = Otherwise
ICF-SNF	1 = Type of certification is Medicare & ICF, SNF & ICF, or Medicare, SNF, & ICF 0 = Otherwise
PROFIT	1 = Individual, partnership, or corporate ownership 0 = Otherwise
CHAIN	If the home is not government owned, this variable indicates whether facility is a member of a group of homes operating under one general authority or general ownership 1 = Yes 0 = No
PCONTRACT	Percentage of operating costs met by contracting with outside group

PONEB	Percentage of rooms with one bed
PTHREEB	Percentage of rooms with three beds
AGE	Average age of patients in the home
PFEM	Percentage of female residents on night prior to the survey
BEDPD	Percentage of patients who need assistance on moving in and out of a bed or chair
BEDD	Percentage of patients who are bedfast
PNOWDIG1	Percentage of depressed patients
PNOWDIG2	Percentage of agitated and nervous patients
PNOWDIG3	Percentage of abusive and aggressive patients
PNOWDIG4	Percentage of confused and senile patients
PNOWDIG5	Percentage of patients with disturbed sleep
PNOWDIG6	Percentage of patients with diagnosis: hardening of the arteries
PNOWDIG7	Percentage of patients with diagnosis: stroke
PNOWDIG8	Percentage of patients with diagnosis: accidents, poisonings, and violence
PNOWDIG9	Percentage of patients with diagnosis: endocrine, nutritional, and metabolic diseases
DEATH	The number of patient deaths in 1972 divided by the census on the night prior to survey
PM1	1 = Type of payment mechanism was cost related (either reasonable, actual, or retrospective) 0 = Otherwise
PM2	1 = Type of payment mechanism was flat rate (incl. pts.) 0 = Otherwise
LIM	Indicates presence of absolute limits on administratiave salaries and/or payments 1 = Yes 0 = No

OCCUPEN	Indicates presence of occupancy penalties 1 = Yes 0 = No
UNIF	Indicates the requirement of maintaing a uniform chart of accounts 1 = Yes 0 = No
LSC	Life Safety Codes were in effect in 1972 1 = Yes 0 = No
MA	Indicates time of patient assessment relative to admission 1 = Pre 0 = Post
CON	Indicates status of the Certificate of Need Program 1 = Functioning 1972 0 = Not functioning
STAFFA	Indicates the requirement that an RN or LPN be present irrespective of the size of the home 1 = Yes 0 = No
STAFFB	Number of nurses required per (SNF/ICF adjusted) patient day for the day shift
SMSA	1 = Not located in SMSA 0 = Located in SMSA
WAGE	1973 average retail wage in the nursing home's county (taken from Area Resource File)
PCI	Percapita income in county in which facility is located
CONC	Measure of concentration of nursing home beds in county, equal to ratio of nursing home beds in facility divided by the nursing home beds in county
BEDPOP	Ratio of nursing home beds in county to population in county

Table T-8.3

Structural Equations
(standard errors in parentheses)

Independent Variable	Dependent Variable: AOC	NHRPD	INDEX	OCCEN	PRIVRATE	PCTPRIV
NHRPD	3.29(.180)***	--	-1.36(.485)**	7.82(1.77)***	12.3(2.15)***	14.7(3.90)***
INDEX	.543(.0997)***	-.218(.144)	--	.658(.765)	3.58(.720)***	4.51(2.34)*
OCCEN	-.0274(.0160)*	--	--	--	.536(.0927)***	--
PRIVRATE	. --	--	--	-.565(.235)**	--	.345(.482)
PCTPRIV	-.0201(.00615)***	.0224(.0150)	-.0123(.0174)	-.0679(.0462)	--	--
BEDS	.00781(.00224)***	.000992(.00077)	.00447(.00067)***	-.0131(.00520)**	-.00350(.00370)	-.0211(.0144)
CHANGE	-1.98(1.13)*	--	--	-20.3(2.47)***	--	--
ADMISPD	185.8(42.7)***	96.9(8.35)***	154.4(46.0)***	-1420.7(160.0)***	-833.0(184.2)***	-1142.1(375.8)***
ICF	-.0111(.447)	-.600(.291)**	-2.07(.217)***	.515(1.49)	8.56(2.03)***	6.63(3.80)*
ICFSNF	.212(.383)	.0311(.182)	-.530(.178)***	.144(.839)	1.40(.554)**	-5.98(1.94)***
PROFIT	-1.36(.444)***	.0906(.151)	-.296(.182)	1.90(1.05)*	1.94(.454)***	-7.10(2.42)***
CHAIN	-.238(.304)	.0363(.0528)	.0980(.0946)	.280(.581)	-.122(.328)	-.642(1.54)
PCONTRCT	.0792(.0361)**	.00312(.00704)	.0224(.0117)*	--	.0880(.041)**	-.190(.203)
PONEB	.0500(.0101)***	-.0116(.005)**	-.00294(.00748)	.00952(.0269)	.0412(.0151)***	.421(.0493)***
PTHREEB	-.0311(.0175)*	.00906(.00377)**	.0128(.00663)*	-.0872(.0382)**	-.0600(.0202)***	-.321(.0925)***
AGE	-.0165(.0353)	-.0124(.00951)	.00239(.0143)	--	.0243(.0346)	--
PFEM	-.0115(.0122)	.000309(.00503)	.0183(.00542)***	--	-.0564(.0196)***	--
BEDPD	-.00194(.00817)	.00334(.00246)	.0117(.00317)***	--	-.0589(.0150)***	--
BEDD	-.0132(.0172)	.00324(.00400)	-.00540(.00610)	--	-.0373(.0184)**	--
PNOWDIG1	.0432(.0184)**	.000930(.00434)	-.0110(.00615)*	--	.0244(.0189)	--
PNOWDIG2	.0541(.0219)**	.00125(.00566)	-.0142(.00740)*	--	.0818(.0239)***	--
PNOWDIG3	.0409(.0195)**	.00341(.00398)	-.00441(.00659)	--	.0244(.0198)	--
PNOWDIG4	.0318(.0170)*	.00627(.00473)	-.00637(.00715)	--	-.0158(.0192)	--
PNOWDIG5	.00773(.0267)	.00205(.00648)	-.0154(.00900)*	--	.0188(.0274)	--
PNOWDIG6	.0191(.0248)	.00252(.00431)	-.00169(.00794)	--	.0182(.0255)	--
PNOWDIG7	.00302(.0188)	-.00152(.00409)	-.0133(.00582)**	--	.0518(.0208)**	--
PNOWDIG8	.0263(.0232)	.000638(.00554)	-.0149(.00765)*	--	.480(.0248)*	--
PNOWDIG9	.0158(.0270)	.000619(.00696)	-.0196(.00912)**	--	.0586(.0296)**	--
PNOWDIG10	.0646(.0334)*	-.00259(.00634)	-.0156(.0103)	--	.0472(.0341)	--
PNOWDIG11	.0430(.0339)	.00128(.00576)	.00208(.0104)	--	-.0278(.0334)	--
PNOWDIG12	.0125(.0233)	.00721(.00553)	-.00334(.00872)	--	.0191(.0232)	--
PNOWDIG13	.0752(.0356)**	-.00287(.00717)	-.0189(.0112)*	--	.139(.0412)***	--
DEATH	2.52(.809)***	-.233(.194)	.185(.302)	--	6.60(1.08)***	--
PM1	-1.74(.456)***	-.200(.0760)***	-.415(.143)***	--	1.39(.521)***	4.32(2.31)*
PM2	-.497(.485)	-.132(.125)	-.601(.150)***	--	2.24(.783)***	.452(2.65)
LIM	-.894(.512)*	-.239(.133)*	.0808(.207)	--	1.17(.563)**	7.08(2.49)***
OCCUPEN	.473(.460)	.336(.157)**	.248(.244)	-.532(.946)	--	-14.1(2.17)***
UNIP	2.04(.415)***	.0635(.125)	.503(.123)***	--	-.977(.650)	--
LSC	-.499(.367)	.175(.112)	.648(.127)***	.474(.748)	-3.46(.726)***	--
MA	-.280(.448)	--	--	--	--	.110(2.20)
CON	2.43(.449)***	--	--	2.464(.847)***	--	--
STAFFA	2.37(.554)***	-.123(.125)	--	--	--	--
STAFFB	2.99(1.63)*	.457(.322)	--	--	--	--
SMSA	-1.06(.354)***	.00616(.113)	-.408(.124)***	-.337(.757)	-.938(.430)**	-.779(2.06)
WAGE	-.279(.466)	.107(.0835)	.323(.148)**	--	-3.12(.568)***	--
PCI	-.000105(.00029)	.0000392(.00005)	.000120(.00009)	-.000107(.00051)	--	-.00122(.00133)
CONC	-.339(.528)	--	.220(.195)	1.67(1.10)	--	-5.94(2.88)**
BEDPOP	-.000592(.00877)	.000490(.00194)	.00604(.00276)**	-.0297(.0150)**	--	.0404(.0403)
(Constant)	5.51	2.37	4.30	86.1	-69.0	-2.54
R_2^2 Adj.	.55751	.33423	.41991	.21772	.34772	.23934
R^2 .	.53824	.30907	.39854	.20357	.32494	.22489
Std. Error	4.84636	.82114	1.48724	9.25572	4.76922	24.38839
Residual Sum of Squares	25342.74263	731.57925	2402.08890	94749.12182	24747.10084	657247.0416

* significant at .10 level
** significant at .05 level
*** significant at .01 level

(1) nursing hours per patient-day (NHRPD),

(2) an index of rehabilitative-type services (INDEX),

(3) average percentage occupancy (OCCEN),

(4) the percentage of total patient-days provided to private-pay patients (PCTPRIV), and

(5) the average price charged to private patients (PRIVRATE).

(These and other variables were defined in Table T-8.2.) Examine now the results for each equation.

T-8.2.1 The Average Operating Cost Equation

The Average Operating Cost Equation relationship is similar to the cost functions estimated in single-equation cost modeling from the same NCHS data in Technical Paper T-4. However, as already noted in Section T-8.1, the single equation cost models did not include the NHRPD variable. In addition, other specifications have been changed, primarily to satisfy requirements for identification of the other structural relations. The numerous service variables are now excluded except as they were used to form the INDEX variable. It was not possible to identify and estimate so many structural relations. In addition, for the sake of economy and to facilitate interpretation, only two of the twelve ADL (activities for daily living) variables were used--BEDPD and BEDD. In a broad sense, these correspond to ambulatory status: BEDPD is partially dependent with respect to getting into and out of bed and BEDD is totally dependent with respect to bed status. Finally, occupancy and bed-size were not indicated in "piece-wise" fashion, as they were in the single-equation model--neither economies nor diseconomies had been found with respect to occupancy and bed-size (see Chapter 3).

Although the reader will no doubt wish to compare this structural cost model with the single-equation results, it is more fruitful to focus on the differences between this structural model and the reduced-form cost equation in the next section. The specifications and purposes of this AOC cost model are fundamentally different from those in the single-equation estimation. In particular, the nursing intensity variable (NHRPD) had been excluded from the single-equation model because of a concern that it presented a serious endogeneity problem. As will be seen from the NHPRD relation, that hypothesis receives strong support. Inasmuch as this variable is

also seen below to be highly important to the structural AOC relation, it should not be surprising that some findings from this analysis are different from those in single-equation estimation. The reduced-form equation should be more nearly comparable to the earlier single-equation estimation.

Recall that although this structural AOC relation is not formally part of the simultaneous-equations system, it nevertheless includes both endogenous and exogenous variables. The so-called endogenous variables are those that have been hypothesized to also depend upon reimbursement and regulatory criteria, whereas the exogenous variables are essentially assumed to be independent or not determined by operators in the short-run. We begin by discussing the results with respect to the endogenous variables--that is, those variables for which we also model the impact of reimbursement and regulation. These variables are excluded in the reduced-form estimation.

All of the endogenous variables included in this structural cost equation are significant at the .10 level or better. As suggested above, the NHRPD (nursing hours per patient day) variable is highly significant with a t-statistic--the coefficient divided by its standard error--in excess of 18. The quantitative importance of this variable is also large; the beta coefficient is 0.455. The significant (negative) coefficient on percentage occupancy (OCCEN) contrasts sharply with the altogether insignificant findings for occupancy variables in the single-equation estimation (Technical Paper #4). It is further interesting that, ceteris paribus, a higher percentage of private patients (PCTPRIV) is now found to be associated with a significantly lower cost level. We attribute both of these results to inclusion of the NHRPD variable and its effect on the public-private mix and occupancy equations. In those equations we find that increased hours of nursing also increases the relative share of beds occupied by private patients as well as the occupancy rate. The point here is that it is nursing hours which create costs but that these costs are offset by having more patients to distribute them across--hence the positive occupancy coefficient--as well as more private patients. The INDEX service intensity variable has the anticipated positive (significant) effect on average operating costs--i.e., the more services, the more costly.

Examine now the findings with respect to exogenous variables.[1] The coefficient of the bed-size variable (BEDS)--our proxy for quantity or scale--is positive and significant at a high level. Although this finding is also at variance with insignificant findings from single-equation estimation, it probably reflects the much less perfect specification of service intensity differences. (The endogenous INDEX variable is the only service intensity measure in this relation, whereas 13 discrete service variables had been indicated in single-equation estimation. Recall that this is not done here for sake of identification.[2]) The ADMISPD coefficient is now halved vis-a-vis the single-equation estimate, and the formerly very significant ICF variable is now completely insignificant. These differences are also probably due to inclusion of the NHRPD variable. That is, the higher costs formerly estimated with respect to patient turnover and SNF-level care no doubt substantially reflect greater nursing intensity, now reflected by the NHRPD variable. This analysis also finds, unlike the earlier model, that cost levels are significantly higher in facilities with more one-bed rooms (PONEB) and significantly lower in facilities with more three-bed rooms (PTHREEB), compared to facilities having two-bed rooms. The diagnostic variables (PNOWDIG1-PNOWDIG13) tend to be more significant (i.e., more variables are significant) than they were in single-equation estimation. Again, this may be due to less complete specification of the service variables. Such potentially spurious correlation could also explain the now significant, positive coefficient obtained for the DEATH variable.

The results obtained for reimbursement and regulatory variables are especially interesting. A somewhat larger negative coefficient, equal to a minus $1.74 per day, is obtained for the PM1 variable, retrospective rate setting compared to prospective. However, the coefficient for the PM2 variable, flat rate relative to prospective, is now insignificant and less

---

[1] Results not discussed are either insignificant or similar to those from single-equation estimation.

[2] Whereas earlier studies had shown cost to be significantly related to bed-size, we likewise believe those findings to result from failure to control completely for relevant service differences. The single-equation cost model in Chapter 3 is probably the first to include relatively complete service measures.

than half of its former magnitude, equal to a minus $0.50 per day. Furthermore, the LIM, UNIF, CON, STAFFA and STAFFB variables are now significant and have the predicted signs. However, the formerly significant OCCUPEN variable is not significant. We suspect that these differences arise because region variables were omitted in this cost equation--i.e., the formerly significant region variables were in part proxying regional differences in reimbursement-regulatory activity.

The SMSA coefficient is now significant, but still negative. However, more importantly, the coefficient estimated for the regional wage level (WAGE) is now highly insignificant. We suspect that this counterintuitive finding is due to multiicollinearity. The WAGE variable is highly correlated with the PCI, CONC and BEDPOP variables, none of which were included in the earlier estimation.

T-8.2.2 The Two Service Intensity Equations

The nursing hours per patient-day (NHRPD) and rehabilitative-type service index (INDEX) variables have been assumed to measure input dimensions of nursing home service quality. Because of the similarity in purpose and specification we consider both relationships at once. Both were estimated by two-stage-least-squares. The results clearly substantiate the thesis underlying the entire multi-equation effort. In particular, the results show that reimbursement parameters are indeed statistically significant determinants of service intensity or "quality" in nursing homes. For example, retrospective reimbursement is found to imply 0.2 reduction in nursing hours per patient-day compared to prospective reimbursement. (The mean value of NHRPD is 2.45.) Similarly, negative impacts on service intensity are found to be associated with flat rate reimbursement. Furthermore, these findings are in accord with those from the cost relation. Whereas retrospective and flat rate reimbursement were there shown to decrease costs, we now find that they also decrease services--and thereupon perhaps quality. That is, cost reductions have been achieved, at least in part, by reducing services--an unintended outcome from the perspective of the rate setter.

The results are now discussed in detail, beginning with the endogenous variables--i.e., those variables that also depend in some degree upon the reimbursement and other variables.

The hypothesized production-possibilities-tradeoffs between service intensity dimensions is supported by the finding of an inverse relationship between them in each of the two service intensity equation. However, the finding is only statistically significant at a high level in the INDEX model. This may be due to the somewhat imperfect definition fo the INDEX service intensity variable as discussed below. The percentage private (PCTPRIV) endogenous variable is not significant at a high level in either relationship. Nevertheless, the results indicate that private patients have a relatively greater preference for nursing intensity and a lesser preference for rehabilitative-type services.

Turning to the exogenous variables, the bed-size scale variable (BEDS) does not indicate any economies of scale with respect to nursing intensity. The coefficient is positive, albeit not significant. The BEDS coefficient is positive and highly significant in the INDEX relation. This is merely an artifact of variable definition. The INDEX variable is merely a measure of service availability and does not necessarily indicate anything about the per-patient intensity of service utilization. A larger facility has more services because problems of indivisibility become less severe--i.e., the fixed costs of services availability, as distinguished from the variable costs associated with service use, can be distributed over a longer patient population and require a smaller increment to the rate.

The ADMISPD results from both service intensity relations are consistent with our expectation that facilities having shorter-stays--usually for rehabilitative reasons--require a greater service intensity. Furthermore, as required for SNF certification, the results indicate that SNF's have a significantly greater nursing and other service intensity.

Continuing the discussion of the exogenous variables, the percentage contracted variable (PCONTRACT) is positive in both the NHRPD and INDEX models, but is significant only in the latter. In both cases, patients in one-bed and two-bed rooms are estimated to require a lesser service intensity. This could conceivably reflect capital-labor substitution--e.g., those with less privacy require more attention and services. On the other hand, it could have something to do with patient placement--e.g., the sickest patients tend to be put into larger rooms. Neither model found that bedridden

patients (BEDD) entailed significant service differentials, as most persons would predict. If this result is correct, it may well reflect the unfortunate consequence of failure to reimburse on a patient-related basis.[1] The government does not recognize such differences in setting the rates, and the nursing homes, in consequence, do not have resources sufficient to meet the great service "needs" of such patients. The coefficients are positive for patients requiring assistance in getting up from a chair (BEDPD), but only the BEDPD coefficient in the INDEX relation is statistically significant.[2] Although some of the diagnostic variables were significant in the INDEX relation, none of them were significant in the NHRPD relation. All of the significant diagnostic variables in the INDEX relation have negative coefficients, perhaps because they indicate patients so incapacitated that they could not benefit from or even use the rehabilitative-type services. The lack of relation to nursing intensity is, however, more surprising. It could also reflect failure to reimburse on a patient-related basis. That is, a constant level of nursing services, as well as rehabilitative services, is provided which does not vary with patient condition, for there is no incentive in reimbursement schemes to provide more nursing or rehabilitative services for "sicker" patients.

Now come the reimbursement and regulatory variables. The results for PM1 indicate that retrospective reimbursement is associated with significantly lower levels of service intensity--NHRPD and INDEX both--relative to prospective reimbursement. An even larger negative effect on INDEX is estimated with respect to flat rate reimbursement. The somewhat smaller effect on NHRPD is not significant. The implications of these results were discussed above. A limit or ceiling on reimbursement (LIM) is found to be associated with a significantly lower nursing intensity--as one would predict for any measure that limits reimbursement and thereupon the resources available to pay for service intensity. Our expectation had been the same for the occupancy penalty variable (OCCUPEN), inasmuch as it is also a constraint on

---

[1] As with the single-equation cost model, a caveat is appropriate here, for the patient debility measures were based on a sample of patients. Hence there is an error in variables problem which results in a negative bias (toward zero) in the estimated coefficients of the patient debility variables.

[2] This finding supports the hypothesis of Skinner and Yett (1973) debilitated patients require more services and nursing care than totally debilitated patients.

nursing home activity and, if it is binding, any such constraint can be assumed to increase costs, or the operator would already have satisfied the constraint in the interest of maximizing profit. However, we instead find that this variable is associated with greater service intensity, and that the relationship is significant in the NHRPD equation. The following is a possible explanation. Looking ahead to the occupancy (OCCEN) demand equation, we see that NHRPD is a highly significant determinant of demand for care within a particular home and therefore of the occupancy level. Thus, it may be that this result arises because facilities have increased their nursing intensity in order to get occupancy up above the penalty level.

A uniform chart of accounts (UNIF) is associated with a significantly higher INDEX value; however, this may merely reflect the greater definitional precision associated with such more formal accounting. That is, an otherwise similar facility not subject to uniform accounting and definitions might not report that it provides so many services because services are not organized and accounted in the same fashion. The medical assessment (MA) and certificate-of-need (CON) variables were excluded for lack of relevance to service intensity, and to achieve identification. The staffing variables (STAFFA and STAFFB) were excluded from the INDEX relation inasmuch as any effect would be mediated by the NHRPD variable already included. The coefficient obtained for the STAFFB variable in the NHRPD relation is positive, as had been expected, but is not quite significant at the usual .10 level.

Neither SMSA or WAGE is significant in the NHRPD relation. This is unexpected and cannot be explained unless nursing wages are generally treated as a pass-through. The findings from the INDEX relation with respect to these variables are entirely perplexing. The coefficient for SMSA is negative and significant, as it should be if it were proxying the higher wage levels in urban areas. However, the coefficient obtained for the regional wage variable (WAGE) is positive and significant. Although per-capita income (PCI) has also been included in the model, it is possible that the wage variable is still proxying regional affluence and that a positive income elasticity of demand for services is more than offsetting a higher input price effect. Of remaining variables, only the county-wide bed-to-population ratio (BEDPOP) is ever significant. We view this variable as a proxy for the relative supply of nursing beds and interpret the positive, significant

coefficient on BEDPOP in the INDEX relation to mean that facilities confronting greater competition tend to offer more services in order to attract patients. The nursing home concentration variable (CONC) was included in the INDEX relation but not in the NHRPD relation. We had reasoned that, if in the extreme, all nursing home beds in a given county or catchment area were concentrated in a single facility, it would be necessary to provide a broader range of services, whereas individual facilities in a given market area might otherwise specialize to some extent. Although the CONC obtains a positive coefficient, as predicted, it is not significant.

T-8.2.3 The Demand Equation (In Occupancy Form)

The percentage occupancy (OCCEN) relation was meant to reflect the market area demand for care within a particular nursing home facility and characteristics of patients within the facility were omitted for this reason. We have assumed that occupancy is not a direct function of the reimbursement rate and the PM1, PM2 and other reimbursement variables have been excluded from the relation for this reason. However, the occupancy penalty variable (OCCUPEN) was retained for obvious reasons--i.e., we hypothesize that facilities seek to avoid the penalty by increasing occupancy. This hypothesis is not supported and may suggest that occupancy penalties are either not enforced or sufficiently stringent. The two most important variables in the model--in terms of their beta coefficients--are the ADMISPD admissions per patient-day and NHRPD nursing intensity variables. The latter service variable is also endogenous and has been seen to be significantly related to the other reimbursement variables. Thus, indirectly, occupancy is also a function of other reimbursement parameters--as is also shown in the reduced-form relation. That is, reimbursement affects the nursing intensity, and the nursing intensity affects occupancy. Another finding of general interest is the result that states with certificate-of-need programs have significantly higher occupancy levels. That is, we find indirect evidence that the CON program is having the desired effect of increasing facility utilization. The result indicates that a CON program increases the occupancy rate by 2.5 percentage points.

Beginning with the endogenous variable, we now consider the results in somewhat more detail. As predicted, there is a significant,

negative coefficient for the private price (PRIVRATE) variable--i.e., private demand is seen to increase as price decreases. Apparently, nursing services are much demanded as shown by the positive and highly significant coefficient obtained for NHRPD. (The beta coefficient is .43.) The INDEX service intensity coefficient is also positive albeit not significant. Private patients apparently prefer a lower occupancy level--or, equivalently, require greater access--as shown by the negative but not quite significant coefficient obtained for percentage private (PCTPRIV).

Examining now the exogenous variables, the BEDS coefficient is both negative and significant, suggesting that the desired occupancy level declines with facility scale--i.e., the same level of access can be achieved at higher occupancy in a larger facility. The CHANGE variable was included merely to control for the short-run disequilibrium associated with changes in facility size; the coefficient is negative and significant. Since most bed-size changes involve incresing the number of beds, this result merely indicates that it takes some time for occupancy to rise to its long-run level. As might be expected, those patients with short stays are willing to pay for greater access and thereupon demand a lower occupancy level as demonstrated by the significant, negative coefficient on ADMISPD. (The beta coefficient is .65.)

It is noteworthy that the desired occupancy level does not vary significantly, ceteris paribus, with certification (i.e., SNF/ICF). This suggests that any occupancy penalties should be applied uniformly across certification levels if other relevant differences have already been taken account of in rate setting (e.g., by using a formula). However, the desired occupancy level is estimated to be significantly higher in for-profit (PROFIT) homes. This may reflect consumer preference for certain unobserved differences between for-profit and non-profit facilities. On the other hand, even though we have sought to specify this OCCEN model as a demand-type relation, it is still possible that this result reflects supply behavior--e.g., the greater efficiency of for-profit facilities in maintaining occupancy. The negative, significant coefficient obtained for the percentage three-bed room variable (PTHREEB) suggests that consumers are willing to pay less for such accommodations, relative to staying in two-bed rooms.

The occupancy penalty (OCCUPEN) is neither positive as had been anticipated nor significant. However, it is possible that the effect of this variable is mediated via the NHRPD variable, as suggested in discussion of the service intensity models. However, this hypothesis is not supported by reduced-form estimation (see Section T-8.3) in which the NHRPD variable is excluded. The implication may be that occupancy penalties are generally ineffectual--or, at least they were ineffectual at the time of the 1972-1973 NCHS survey. However, regulation has not been altogether ineffective in maintaining occupancy. The positive and highly significant coefficient obtained for the CON variable indicates that certificate-of-need programs have raised occupancy levels. Furthermore, this result obtains, even after controlling for the county-wide bed-to-population ratio (BEDPOP) and consequently underestimates the total effect of CON programs, because the mechanism by which the CON program works is to decrease or limit growth in the aggregate number of nursing home beds. As might be expected, a higher bed-to-population ratio is seen to mean a significantly lower average occupancy. Although the coefficient on the CONC variable is not quite significant, its positive coefficient is consistent with the hypothesis that occupancy levels are higher in markets with fewer competitors.

T-8.2.4 The Inverse Supply Function

An inverse supply function is a function in which, instead of having the quantity supplied as a function of the supply price and other variables, the supply price is a function of quantity supplied and those other variables. (The two are formally equivalent.) This relation models the private price of care (PRIVRATE) and relates the price at which nursing home output will be supplied by a nursing home, given the regulatory and other constraints on its behavior, to characteristics of the output, input prices, etc. The results further substantiate the importance of simultaneity in understanding nursing home behavior. All three endogenous variables are significant at the .01 level or better. For example, the private rate is now found to be a significant function of occupancy whereas occupancy had been shown above to be a significant function of the private price. The results obtained with respect to reimbursement variables are especially interesting. The PM1, PM2 and LIM (retrospective rate, flat rate, and absolute rate

ceiling) coefficients are all positive and significant. In particular, the PM1 (retrospective reimbursement) coefficient is almost equal in magnitude but opposite in sign to the PM1 coefficient obtained in the AOC equation. Although it is premature to draw any firm conclusion, these results strongly suggest that cost containment measures with respect to public patients have exerted upward pressure on the prices charged to private patients--that, at least in some degree, retrospective and flat rate reimbursement (relative to prospective reimbursement) and reimbursement ceilings have led to private patients subsidizing public patients. That is, those reimbursement parameters that have been shown to lower average costs (as defined by public accounting standards) are now seen to increase the rates charged to private patients. Whatever the mechanism, this surely has never been the intent of rate setters and raises serious questions about the equity consequences of reimbursement policy.

We now discuss other results. The coefficient obtained for the percentage occupancy (OCCEN) endogenous variable is positive and significant. This variable expresses the bed-day output level in relation to capacity; the findings indicate that a greater quantity-supplied entails a higher supply price--i.e., an upward sloping supply relationship. The supply price increments associated with service intensity--both NHRPD and INDEX--are four or more times the size of the estimated cost increments in the AOC relation. This finding would appear to argue in favor of cost-related reimbursement versus some free market-type approach. That is, it appears that private patients are being charged considerably more for these services than the (average) marginal cost of providing them.

Consider the exogenous variables. The rather large negative and significant coefficient obtained for the ADMISPD variable is unexpected. Multi-collinearity appears to be the problem since the significant coefficient obtained for the ICF variable is also reversed from what it should be.[1] The PROFIT coefficient is positive and significant, compared to a negative and similarly significant coefficient in tne cost relation. This tends to suggest that private patients in for-profit facilities are subsidizing the care of public patients more so than private patients in nonprofit facilities. Either the "true" full cost is not being reimbursed for public patients or

---

[1] Although the simple correlation between ADMISPD and ICF is only -0.18, that does not preclude the possibility of much more serious collinearity once other variables have been taken account of.

the private homes are earning excess profit on the private patients. The PCONTRACT result implies that contracting is more expensive, contrary to expectation. It is possible that facilities are contracting for unobserved amenity-type services. On the other hand, the much-alleged "kick-backs" associated with fraudulent contracting for purposes of gaining greater public reimbursement may also be reflected in the private rate to avoid suspicion. The PONEB and PTHREEB bed-room distribution results are consistent with expectations--namely, that it is cheaper to provide services when you have more patients in the same room. The highly significant, negative coefficient on PFEM is not consistent with earlier indications that it is more expensive to take care of women. This result probably reflects less complete control for debility (only two ADL's) in this multi-equation analysis. It is well known that women have better health than similarly aged men and therefore should cost less to care for. However, that is only part of the explanation. The reader will find from the reduced-form equation in Section T-8.3 that that PFEM coefficient is reversed in sign and still significant. We attribute this to the greater "demand" of females for service intensity, especially rehabilitative-type services, and the consequent impact of service intensity in cost. That is, the reduced-form result finding is consistent with the single-equation cost estimation. The BEDPD and BEDD disability variables are both significant but have unexpected negative signs. However, as we shall see, the predicted positive (significant) coefficient is obtained for BEDPD in reduced-form estimation; the BEDD coefficient is insignificant. If the more disabled patients tend to be public patients, the results here may reflect a reduced willingness of private patients to pay for care in facilities having "sicker" patients. (Unfortunately, our data cannot adequately discriminate between the characteristics of public and private patients.) We do not attempt to interpret the coefficients obtained for the diagnostic variables. They were merely included to provide more complete control for relevant patient differences. The finding for DEATH is consistent with the result in the AOC relation.

We had anticipated that PRIVRATE would be higher in urban areas and in higher wage counties. Those expectations were not borne out. In fact, the results were the opposite. We cannot explain this finding except

to suggest that more urban and more affluent areas may be experiencing more price competition. Explicit measures of competition were unfortunately not included in the model. It will be important in any subsequent modeling to give further attention to this problem.

T-8.2.5 The Relative Supply Function

The relative supply or access of nursing home care to public and private patients is modeled using percentage private (PCTPRIV) as the dependent variable. Here again, we find that rate setting has had an impact on a variable of interest besides cost. In particular, those features of rate setting that were earlier shown to confine costs--i.e., have negative coefficients in the AOC relation--are now found to reduce the relative access of public patients to nursing home care. Such an effect is easily forecasted. The lower the reimbursement offered for public patients in relation to the price bid by private patients--i.e., the lower the relative price of public patients--the greater will be the relative supply of beds to private patients. Of course, indirectly this has the effect of controlling aggregate reimbursement expense inasmuch as there are fewer publicly-reimbursed patients in nursing homes. The significant, positive coefficients on NHRPD and INDEX indicate that facilities providing a greater service intensity attract more private patients, on the average. The private price (PRIVRATE) has the predicted sign but is not significant.

The negative and significant coefficient on ADMISPD suggests that private patients have longer average stays. The significant, positive coefficient on the ICF variable may simply reflect the lack of Medicare reimbursement for such lower-level care. The results also show that for-profit facilities have significantly fewer private patients and that facilities with fewer patients per room attract relatively more private patients.

Several of the reimbursement variables are significant. In particular, the PM1 coefficient is significant and positive. This result implies that retrospective reimbursement, as compared to prospective reimbursement, reduced the access for public patients. The PM2 coefficient is also positive albeit not significant. Finally, we here find evidence that occupancy penalties (OCCUPEN) have had a beneficial effect for public patients. It may be that facilities have boosted their occupancy levels by taking

public patients and thereby altering the public-private patient mix. It had been thought that medical assessment (MA) would increase the percentage private mix because of the greater cost associated with admitting public patients. The MA coefficient is positive, as predicted, but it is not significant.

The CONC variable is negative and significant. This variable was merely included as a control inasmuch as the concentration of area beds in a lesser number of facilities means that there is necessarily less prospect for facilities to specialize in either public or private patients.

In conclusion the following has been shown: (1) The results affirm the importance of simultaneity issues in analyzing nursing home behavior. It is a more complex behavioral environment than can be portrayed by a single-equation causal nexus. For example, it has been found that an increase in service intensity increases operating costs, but also that it increases occupancy and percentage private patient mix, both of which lower costs, ceteris paribus. In addition, a greater service intensity implies a higher private rate, which in turn implies a lower occupancy level and hereon higher operating costs. On the other hand, a lower occupancy level implies again a higher private rate. Whereas this analysis has not investigated all potential inter-relationships of this kind, it has at least shown that they are important and relevant to rate setting.

(2) The reimbursement and regulatory variables do indeed impact variables of interest other than cost. In particular, we have shown from this structural analysis that reimbursement variables also affect nursing and other service intensity, the rates charged to private patients and the private-public patient mix--a proxy for access. In addition, occupancy levels are impacted indirectly. At the same time, we are reluctant to make strong policy inference from our findings. Whereas many of the results are consistent with informal hypotheses, others are counterintuitive and may suggest either mis-specification or model instability across states. We suspect that the latter has been the problem, especially given that the states' analyses have already given evidence for that view. Apparently, more variables are needed to parameterize relevant state-to-state differences in reimbursement policy.[1] In any event, we freely acknowledge that

---

[1] A more fruitful alternative in the near-term future may be estimation of a similar multi-equation system from data for a single state (e.g., New York).

sufficient attention has not yet been given to specification of the structural relationships. Our effort should be viewed as exploratory. Whereas it has been shown above to offer a useful framework for understanding nursing home behavior, alternative frameworks should similarly be developed and estimated. In particular, we suspect that estimation using a catchment area model also would be fruitful in contrast to our facility level analysis. That is, another data base may be more appropriate.

T-8.3 The Reduced-Form Equations

The reduced-form equations were estimated by ordinary-least-squares and are reported in Table T-8.4. Formally, these equations are derived from the structural equations by solving for each of the endogenous variables so as to eliminate other endogenous variables from the right-hand side. This is accomplished in much the same way that n linear relations are used to solve for n unknowns, except that the unknowns remain a function of the exogenous variables. The reduced-form equations generally give the more policy-relevant information because reduced-form coefficients give the equilibrium impact of a change in any exogenous variable on any endogenous variable. For example, the PM1 coefficient in the AOC relation reflects not only the direct impact on cost but also its effect on the other endogenous variables and their effect on cost. That is, it reflects not only the direct impact of PM1 on average operating cost, but also its impacts on NHRPD, INDEX, PRIVRATE and PCTPRIV and their impact on cost. Likewise, the NHRPD relation reflects not only the direct effect of PM1 on nursing intensity but also its effect on other endogenous variables and their "feedback" effect on nursing intensity.

It seems likely that the rate setter will wish to "short-circuit" some of the structural effects and, for example, maintain service intensity thresholds by regulation or reimbursement incentive much as has been done with occupancy penalties. That is, the policy alternatives being considered in this study are not limited to the reimbursement and regulatory programs parameterized in the econometric work. Thus, we are not merely interested in the final effects but also in the mechanism by which those effects obtain. However, because of such "feedback" effects, it is difficult to gauge whether or not reduced-form results are consistent with expectations. This is more readily done in the context of the structural relations above. In fact, the

Table T-8.4

Reduced Form Equations
(standard errors in parentheses)

Independent Variable	Dependent Variable: AOC	NHRPD	INDEX	OCCEN	PRIVRATE	PCTPRIV
BEDS	.0105(.00253)***	-.0000167(.00037)	.00442(.0006)***	-.0156(.00415)***	.00389(.00209)*	.00436(.0108)
CHANGE	.127(1.28)	.632(.186)***	-.356(.337)	-.15.1(2.10)***	-1.47(1.06)	9.72(5.45)*
ADMISPD	525.0(44.3)***	95.2(6.43)***	22.7(11.6)*	-621.5(72.6)***	86.6(36.6)**	153.4(188.1)
ICF	-2.01(.485)***	-.449(.0705)***	-1.35(.127)***	-2.22(.795)***	-2.50(.401)***	-7.36(2.06)***
ICFSNF	-.610(.439)	-.152(.0637)**	-.187(.115)	-.0397(.718)	-.916(.362)**	-10.1(1.86)***
PROFIT	-1.50(.513)***	-.0643(.0746)	-.0788(.135)	1.37(.841)	1.53(.424)***	-7.81(2.18)***
CHAIN	-.0986(.354)	.0445(.0514)	.0259(.0929)	.209(.579)	.497(.292)*	.741(1.50)
PCONTRCT	.0726(.0418)*	-.0101(.00606)*	.0365(.0110)***	-.0876(.0684)	.0430(.0345)	-.220(.177)
PONEB	.0304(.0115)***	-.00403(.00167)**	-.00114(.00302)	-.326(.0188)*	-.0330(.00948)***	.329(.0487)***
PTHREEB	-.0141(.0203)	.00176(.00294)	.0114(.00533)**	-.0311(.0332)	-.0323(.0168)*	-.207(.0861)**
AGE	-.0325(.0409)	-.00107(.00594)	-.00130(.0108)	.0490(.0670)	.0287(.0338)	.518(.174)***
PFEM	.00167(.0141)	.00431(.00204)**	.00865(.00369)**	.00731(.0230)	.0264(.0116)**	.275(.0597)***
BEDPD	.0143(.00943)	.00496(.00137)***	.00373(.00248)	.00302(.0154)	.0193(.00778)**	.100(.040)**
BEDD	-.00783(.0211)	.00169(.00307)	-.00577(.00555)	.0531(.0346)	-.0142(.0175)	-.121(.0898)
PNOWDIG1	.0422(.0214)**	.000757(.00311)	-.0102(.00563)*	.00117(.0351)	-.00352(.0177)	-.101(.0910)
PNOWDIG2	.0500(.0254)**	-.000817(.00369)	-.0109(.00668)	-.0266(.0417)	.0103(.0216)	-.191(.108)*
PNOWDIG3	.0493(.0226)**	.00342(.00328)	-.00857(.00594)	-.0115(.0371)	.0250(.0187)	-.0774(.0960)
PNOWDIG4	.0445(.0197)**	.00487(.00287)**	-.0108(.00518)**	.0174(.0323)	.0107(.0163)	-.151(.0838)*
PNOWDIG5	.00529(.0310)	.000438(.00450)	-.0129(.00813)	.0282(.0507)	-.0106(.0256)	-.193(.132)
PNOWDIG6	.0312(.0288)	.00508(.00418)	-.00807(.00756)	-.0215(.0472)	.0338(.0238)	.0377(.122)
PNOWDIG7	-.00350(.0218)	-.000466(.00317)	-.0119(.00573)**	-.00901(.0357)	-.00633(.0180)	-.0624(.0926)
PNOWDIG8	.0242(.0269)	.000575(.00391)	-.0139(.00707)**	-.0201(.0441)	-.000122(.0222)	-.149(.114)
PNOWDIG9	.00991(.0314)	-.000576(.00456)	-.0159(.00825)*	-.0101(.0514)	-.0120(.0259)	-.216(.133)
PNOWDIG10	.0533(.0389)	-.00180(.00564)	-.0115(.0102)	.0282(.0636)	-.00777(.0321)	-.0805(.165)
PNOWDIG11	.0436(.0395)	-.00103(.00573)	.00527(.0104)	.0107(.0647)	-.0172(.0326)	-.0667(.168)
PNOWDIG12	.0235(.0271)	.00236(.00394)	-.00410(.00712)	-.0328(.0444)	.0131(.0224)	-.230(.115)**
PNOWDIG13	.0633(.0414)	-.00353(.00602)	-.0119(.0109)	-.109(.0679)	-.0111(.0342)	-.158(.176)
DEATH	2.17(.933)**	-.166(.136)	.313(.245)	-5.78(1.53)***	2.56(.771)***	6.58(3.96)*
PM1	-2.31(.529)***	-.0859(.0768)	-.404(.139)***	.548(.866)	-1.67(.0437)***	2.26(2.25)
PM2	-.872(.561)	-.0454(.0815)	-.565(.147)***	-1.05(.919)	-.132(.463)	-2.59(2.38)
LIM	-1.50(.590)**	-.236(.0857)***	.417(.155)***	-1.54(.966)	-1.59(.487)***	5.06(2.50)**
OCCUPEN	1.09(.529)**	.0701(.0768)	.261(.139)*	-1.04(.866)	1.23(.437)***	-10.6(2.25)***
UNIF	2.41(.481)***	.0782(.0699)	.373(.126)***	.0394(.788)	1.85(.397)***	4.10(2.04)**
LSC	.176(.423)	.184(.0614)***	.317(.111)***	1.08(.693)	.213(.349)	3.64(1.80)**
MA	-.596(.514)	-.210(.0747)***	.613(.135)***	-.219(.843)	-.927(.425)**	-1.74(2.18)
CON	2.61(.522)***	.0700(.0758)	-.131(.137)	.761(.855)	3.51(.431)***	-2.10(2.22)
STAFFA	2.51(.642)***	.143(.0932)	-.355(.169)**	.0248(1.05)	2.37(.530)***	6.75(2.73)**
STAFFB	3.15(1.89)*	-.114(.274)	.835(.496)*	7.90(3.10)**	1.43(1.56)	-15.1(8.03)*
SMSA	-1.12(.409)***	.0383(.0594)	-.429(.107)***	.258(.669)	-1.47(.337)***	-2.97(1.74)*
WAGE	.100(.541)	.102(.0786)	.179(.142)	1.70(.887)*	-.276(.447)	.286(2.30)
PCI	.000215(.00033)	.0000178(.00005)	.0000807(.00009)	-.000212(.00054)	.000389(.00027)	-.0000915(.00141)
CONC	-.597(.609)	-.194(.0884)**	.510(.160)***	1.01(.997)	.234(.503)	-6.43(2.58)**
BEDPOP	.00341(.0102)	.000435(.00148)	.00500(.00267)*	-.0207(.0167)	.00166(.00841)	.0356(.0432)
(Constant)	10.3	1.93	1.55	89.9	9.21	-2.64
R^2 Adj.	.373	.311	.402	.205	.354	.251
n_2	1127	1127	1127	1127	1127	1127
R^2	.39718	.33745	.42479	.23526	.37951	.27962
Std. Error	5.64621	.81990	1.48301	9.24802	4.66230	23.97385
Residual Sum of Squares	34525.66316	728.03705	2381.87318	92624.51308	23541.18762	622449.12727

* significant at .10 level
** significant at .05 level
*** significant at .01 level

validity of a reduced-form relation is most reliably tested by evaluating the reasonability of the structural model from which it derives.[1] Finally, although the reduced-form coefficients may give the aggregate effect it will be useful in designing a reimbursement system to also understand the structural process.

T-8.3.1 Facility Variables

Rather than review the specific equations, we review the overall pattern of results.

First of all, we discount the general significance of the bed-size variable (BEDS) because of certain specification problems which were discussed previously. The ICF certification variable is uniformly significant at the .01 level in all six structural relations. ICF certification, relative to SNF certification, implies a lower average operating cost (-$2.01), a lower nursing (-0.45 nursing hours per patient-day) and rehabilitative service intensity, a lower occupancy (-2.2 percentage points), a lower private rate (-$2.50) and a smaller percentage of private patients (-7.4 percentage points). (The negative effects on occupancy probably reflects its negative structural relation to the service intensity variables and this strong positive relation of service intensity to occupancy and percentage private.)

The admissions per patient-day variable (ADMISPD) is found to have significant, positive effects on average cost and the two service intensity variables. These service results support earlier hypotheses that short-stay patients are more expensive because of their interest in being rehabilitated and consequent demand for services. A greater rate of patient turnover is also found to significantly reduce occupancy as one would expect. Furthermore, ADMISPD has the same positive (significant) relation to the private rate as it did to average operating cost. However, its coefficient is much smaller, perhaps because private patients do not involve the fixed paperwork and other expense that public patients do. The estimated cost per admission (equal to the coefficient on ADMISPD) is only $87 compared to $525, whereas the increment to the private rate is only $87.

---

[1] Although it is now common to estimate so-called reduced-form equations without even specifying the underlying structural relations, that is surely a less desirable alternative. On the other hand, if that had been done here, more "complete" and realistic reduced-form relationships could have been estimated since identification of the structural relations would no longer be an issue.

The for-profit (PROFIT) facilities are found to have significantly lower operating costs (-$1.50) but to charge a significantly higher rate to private patients (+$1.53). As noted in discussion of similar indications from the structural relations, this suggests either that private patients are subsidizing public patients in for-profit facilities or that such facilities are earning excess profits on private patients. However, for-profit homes also have significantly fewer private patients (-7.8 percentage points). This is consistent with an earlier hypothesis that for-profit facilities provide fewer amenity-type services.

The CHAIN variable, identifying members of for-profit chains, is significant only in the PRIVRATE equation and implies a higher rate. Although there is no similar finding from the average cost relation, this may reflect a somewhat higher service intensity in chain facilities. (The NHRPD and INDEX coefficients are positive, but not significant.)

Contracting is again shown from the PCONTRCT variable to have a positive (significant) IMPACT on cost. The higher cost associated with contracting may explain contracting's negative effect on nursing intensity; ceteris paribus, there is a production-possibility tradeoff. Both cost money, and if you have more contracting, at a given cost level, you must have less nursing services. A positive and significant coefficient is obtained for the PCONTRCT variable in the INDEX relation. This merely shows that contracted services include some of the rehabilitative ones in rehabilitative-type service index. The bed-room distribution variables (PONEB and PTHREEB) indicate, also consistent with earlier results, that a smaller number of patients per bed-room is more costly. Furthermore, we again find that the service intensity variables vary directly with the number of patients per room. As noted below this could be due either to capital-labor substitution or systematic patient placement. The PRIVRATE results further confirm that private patients are willing to pay a premium for more private rooms, furthermore, they are willing to pay much more than the apparent cost differentials. (Of course, this may merely reflect the exclusion of capital cost from the cost variable.)

T-8.3.2 Patient Variables

The average patient age variable (AGE) is only significant in the percentage private (PCTPRIV) relation and implies that private patients tend

to be older. A possible inference from this result may be that because the cost of care is actually paid by the patient or patient's family, private patients tend to defer entering a nursing home until later in life. If so, lack of co-insurance provisions in public payment programs for nursing home care may imply that over-utilization results. Female patients (PFEM) are seen to require a significantly greater service intensity--e.g., 0.4 additional nursing hours per day. In addition, female patients are charged a higher private rate (+$2.64 per day); the estimated $1.67 increment to cost is not significant. Furthermore, a significantly larger percentage of private-pay patients are female. Whereas two of the coefficients are not quite significant at the .10 level, the partial debility variable (BEDPD) indicates that such debility implies increased service intensity (both NHRPD and INDEX), increased costs and a higher rate charged to private patients. Thus, results are consistent with expectations. In addition, partially disabled patients tend to be private-pay patients. This is consistent with our generic hypothesis above that private patients tend to defer entering a nursing home until they have greater "need." The total debility variable (BEDD) is not significant in any of the reduced-form equations. As noted in discussion of the structural results, this may reflect lack of reimbursement on a patient-related basis and may raise questions about the quality of care received by such patients. The patient diagnostic variables (PNOWDIG1-PNOWDIG13) are most significant in the AOC and INDEX relations. Patients with diagnosed illnesses are found to cost more but demand fewer rehabilitative services, perhaps because they are too incapacitated to use or benefit from them. The DEATH variable may also proxy patient well-being and is found to be positively related to cost and the private rate. The greater patient turnover associated with death and the private rate. Since service intensity had been found from the structural estimation to increase occupancy, we would have thought a positive reduced-form coefficient would obtain. One explanation might be that the occupancy variable is itself endogenous in the sense that occupancy variables are only adopted in states having had historically low occupancy levels. The highly significant and negative coefficient on the PCTPRIV variable is consistent with that view. That is, that result suggests that occupancy levels have been increased by taking more public patients and thereby altering the otherwise natural public private-patient mix accounts

for its negative coefficient in the occupancy relation. Furthermore, the positive (significant) coefficient in the PCTPRIV relation is also consistent with the hypothesis that private patients are less well, having deferred entry to a nursing home until physical deterioration is more advanced.

T-8.3.3 Reimbursement and Regulatory Variables

The reimbursement and regulatory variables have the greatest interest in this study. We see that retrospective (PM1) and flat rate (PM2) reimbursement, relative to prospective reimbursement, lower the average operating cost level by $2.31 and $0.87 per patient-day, respectively. The former estimate is larger than that obtained earlier in single-equation estimation and the latter estimate is somewhat smaller. The PM1 coefficient is larger because prospective reimbursement leads to an increase in rehabilitative service intensity and this increases cost. (Rehabilitative services--here seen to be endogenous--were included in the single-equation model.) The lower PM2 coefficient is due to the greater effect of flat rate reimbursement in the structural model on the private rate variable and its feedback effect on cost. The PM1 and PM2 coefficients are not significant in the NHRPD relation; however, it is noteworthy that both are negative. Furthermore, both variables elicit negative coefficients in the INDEX relation and both are significant at the .01 level or better. Apparently the greater effect of these reimbursement variables on rehabilitative services feeds back into the nursing intensity structural relation and "swamps" the significant structural effect of these reimbursement variables. It is also found that retrospective reimbursement reduces the rate charged to private rate. However, this can be misleading. The private rate is reduced only because such reimbursement reduces the service intensity and this we have seen from the structural relation, if one controls for service intensity, the result is the opposite. A limitation or ceiling (LIM) on reimbursement is found to reduce costs by $1.50 per day, compared to $.89 in the structural relation. It has a similarly negative impact on nursing intensity. We attribute its (significant) positive coefficient in the INDEX relation to feedback from the NHRPD relation, much as was hypothesized above with respect to the PM2 variable except in the opposite direction. Although the negative LIM coefficient in the occupancy relation is not quite significant, we attribute it to LIM's negative effect on NHRPD and thereon its positive relation to occupancy. Its

negative coefficient in the PRIVRATE equation can be similarly explained. Finally, much as we found in the structural relation, an absolute reimbursement ceiling is found to significantly reduce the access of public patients--proxied by PCTPRIV--to nursing home care. Such limitations widen the gap between private and public rates in some facilities and make it relatively more attractive to accept private patients.

The occupancy penalty variable (OCCUPEN) does increase occupancy levels; in fact, the parameter estimate, although insignificant, is negative. It is found that occupancy penalties increase service intensity and thereupon increase operating costs.

Uniform accounting (UNIF) is found to significantly increase both operating costs and the private rate--the one by $2.41 per day and the other by $1.85. Although the magnitudes are comparatively large, we had anticipated that uniform accounting would involve some increment to cost. However, the positive effect on INDEX is altogether unexpected. We had earlier suggested, in discussion of a similar finding from structural estimation, that it was merely an artifact of having services uniformly defined. However, the significant relation now found between PCTPRIV and the UNIF variable suggests that such an interpretation is not correct. It appears that more rehabilitative services are actually being offered and that this greater service intensity attracts more private patients. We cannot explain this and suggest that further attention be given to this issue in subsequent work.

The life and safety code requirement is found to increase both nursing and other service intensity but not to have any significant relation either to cost or the private rate. Perhaps multi-collinearity has been a problem; the AOC and PRIVRATE coefficients are both positive. The positive and significant coefficient in the PCTPRIV relation reflects the impact of greater service intensity on private demand for nursing home care.

Although the medical assessment requirement variable (MA) elicits several significant coefficients, the pattern of results is uninterpretable. We would omit this variable in subsequent work. It has no significant relation to cost.

A Certificate-of-Need program was found from structural estimation to increase occupancy levels. The CON variable is not significant, however, in the reduced-form occupancy model, although the coefficient is still positive. This result is also puzzling and may further signal that multi-collinearity has been a severe problem in reduced-form estimation of these reimbursement-regulatory coefficients. The CON variable is found to be associated with positive and significant increments to cost and the private rate.

The staffing requirement variables (STAFFA and STAFFB) were both estimated to increase operating costs and private price levels. However, other results are inconsistent and we must again raise the spectre of multi-collinearity. However, mis-specification is another possibility; in particular, the importance of a STAFFA requirement varies inversely with facility size and STAFFB requirements may vary substantially among states having them.

T-8.3.4 Region Variables

As we found from reduced-form estimation, the SMSA and regional WAGE variables gave unexpected results. We find the same here. Average operating cost levels, private rates and INDEX levels are herein found to be lower in SMSA's. We cannot explain this pattern. The WAGE variable is only significant in the occupancy equation. It may be the region variables omitted in this analysis had been proxying program differences that are systematically correlated with wage levels. The per capita income variable (PCI) is uniformly and highly insignificant. The nursing home concentration (CONC) and nursing bed-to-population (BEDPOP) results are not fundamentally changed from those obtained in reduced-form estimation.

The reader is reminded that we have even less of a "benchmark" for assessing the reasonability of reduced-form findings then we did with respect to the structural results. Nevertheless, certain results and patterns in the results are undeniably suspicious and suggest multi-collinearity and other specification problems. In general, we would be somewhat more comfortable drawing inferences from the structural results, and simulating the feedbacks from the structural relations. Even so, the reduced-form results affirm the essential finding from structural estimation that reimbursement and regulatory activities have influenced nursing home behavior along relevant dimensions in addition to cost.

T-8.4 The Implications For Design of A Reimbursement Scheme

Despite various misgivings about the multi-equation model developed herein, the experience gained to date with such sophisticated modeling nevertheless has led to insights concerning reimbursement issues. In addition, the findings from our admittedly imperfect multi-equation constitute the best evidence available in certain areas. Perhaps the greatest contribution of this multi-equation study will have been merely to dramatize the importance of endogeneity concerns in nursing home reimbursement. A "black box" approach to rate setting is shown to be unsatisfactory. That is, rate setting is shown both conceptually and empirically to have unintended and offer undesired consequences on variables of interest in addition to cost. Whereas "lip service" is often given to such possibilities, this study affirms that such possibilities must be taken seriously and that rational rate setting cannot proceed apart from a more comprehensive understanding of the nursing home behavioral environment. However, we do not go so far as to suggest that the present multi-equation modeling effort will furnish an adequate guide for that purpose. Although it certainly would be useful to continue multi-equation modeling efforts in this area, it would be more fruitful for shorter-term policy interests to promote a dialogue between rate setters and nursing home operators. The naivete of existing reimbursement systems is virtually prima facie evidence that the communications gap is serious. While rate setters no doubt share this perception in some degree, we must conclude, after pursuing the most sophisticated cost analyses to date, that it would be a mistake for them to rely upon statistical analyses to "bridge" the informational gap. Whereas statistical analysis can be useful in illuminating certain partial questions, it cannot in any sense provide a complete and comprehensive guide to reimbursement policy.

This is not to suggest that the government allows the nursing home industry to set rates. Nor is it to suggest that the nursing home industry does not have a vested interest and that spokesmen for the industry would not carefully screen the kinds of information given to rate setters. However, as in any industry, it is always possible to identify certain responsible individuals who have the greater "public interest" at heart and who are willing to speak openly about nursing home problems, reimbursement alternatives and their implications. Until such "dialogue" has taken place, it is

unlikely that the technician can contribute much to development of an "equitable and efficient" reimbursement scheme.

Nevertheless, we believe that much of practical importance has been learned from the multi-equation modeling effort. In particular, the following implications emerge from our findings:

- The demand for nursing home care by private-pay patients is not price inelastic. By not confronting public patients with co-insurance or other cost controlling incentives, public reimbursement of nursing home care biases such care towards over-utilization vis-a-vis less expensive substitutes (e.g., home care). An efficient reimbursement scheme should somehow attempt to circumvent such bias.

- On the other hand, cost-controlling reimbursement features appear to have a negative impact on the access of public patients to nursing home care. Thus, it may be premature to focus on continuing costs; the rates may already be too low to satisfy the existing demand for public care. (Whereas this effect biases toward under-utilization, it does not correct the above bias toward excess utilization by discriminating between those with greater or lesser intensity of preference for nursing home care. It is instead "first come-first served" or queuing.)

- Cost-controlling reimbursement features also imply reductions in service intensity, and by presumption, the quality of care. The rate setter must decide whether or not such reductions in service intensity and quality are desired.

- The lack of strong relation between certain debility-type variables and the service intensity (and cost) of care raises serious questions about the quality of care being given to more debilitated patients. If the quality of care is in fact deficient for such patients, it will be especially important to develop a patient-centered reimbursement scheme that explicitly takes account of differential patient need.

- We have no strong evidence that occupancy penalties on other reimbursement criteria have had any impact on occupancy rates. However, we have some evidence that CON has increased occupancy levels. These finding are generally consistent with the view that the more usual problem is not excess capacity but rather excess demand.

- Private patients are paying higher rates than public patients to begin and cost-controlling reimbursement features increase the differential. If nursing homes are not earning excess profits, the implication seems to be that private patients are subsidizing public patients and the extent of subsidy is increased by effective reimbursement controls. An equitable reimbursement scheme should not require such cross-payor subsidization.

In general, the pattern of results has been such as to call into question the underlying purpose of this study--namely, to design an efficient reimbursement scheme that will contain per-unit cost. That is, our findings, admittedly inconclusive, signal that historic cost levels have been too low to satisfy other social goals in the nursing home area--e.g., quality, access and equity. Given existence of problems in these areas, it seems unlikely that single-equation analysis of historic cost levels can furnish a very dependable guide to developing a reimbursement system that is efficient in a broader sense than merely reducing or limiting per-unit costs. Furthermore, in terms of limiting costs per se it seems that non-rate setting alternatives to cost control have much greater potential for reducing system costs--e.g., (1) enhancing access and thereby reimbursing patients at nursing home rates rather than maintaining them in hospitals pending admission to a nursing home and reimbursing at the higher hospital rate, and (2) incorporating co-insurance or other incentives into reimbursement policies to avoid over-utilization.

BIBLIOGRAPHY

Aigner, D.J. and S.F. Chu. "On Estimating the Industry Production Function." Vol. 58, No. 4, September 1964.

American Health Care Association. Long Term Care Facts. Washington: American Health Care Association.

Applied Management Sciences, Inc. Analysis of Incentive Reimbursement System for Health Care and LTC Services Provided to the Elderly and Long-Term Disabled. Final Report, prepared under contract no. HEW 100-76-0029 for DHEW, Office of the Assistant Secretary for Planning and Evaluation. August 1976.

_____. Report on Systems of Reimbursement for Long-Term Care Services. Contract no. HRA 106-74-0186. April 1977.

Ball, Albert Lowey. "Costs of Long-Term Care Data and Organization." Report to the Massachusetts Department of Public Health. 1975. (Mimeograph.)

Batelle Health Care Study Center. "Information Required for Examination of Rates of Return in Nursing Homes." Unpublished questionnaire data made available by Suresh Malhotra, N.D.

Batelle Memorial Institute. Profit and Growth: Allowances for Owner/Operator Contributions in a Cost Related Prospective Payment for Services Systems for Nursing Home Care. Washington: American Health Care Association, 1975.

Baumgarten, Steven A., Santoch K. Choudhury, Tanniru R. Rao, and L. Winston Ring. "Study of Nursing Homes: The Patient as Consumer." Proceedings of the American Marketing Association Conference. Pps. 452-456. 1975.

Baumgarten, Steven A., Tannirir R. Rao, and L. Winston Ring. "A Descriptive Model of Consumer Choice Processes Among Nursing Home Patients." Advances in Consumer Research. Vol. III. Ed. B. Anderson. (Proceedings of the Association for Consumer Research Sixth Annual Conference.) 1975.

Baumol, William and Allen Klevorich. "Input Choices and Rate of Return Regulation." Bell Journal of Economics. Autumn 1970.

Beattie, Robert and Hernon S. Jordan. "A Preliminary Analysis of a Survey of Massachusetts Hospital Patients who Were Ready for Discharge and Awaiting Placement in Other Facilities on March 20 1974." Research Series, Report No. 1. Office of Health Planning and Statistics. April 11, 1975.

Bicknell, William J. and Diana Chapman Walsh. "Performance Oriented Regulation of Long Term Care: A Proposal."

Bishop, Christine E. Nursing Home and Rest Home Costs in Massachusetts: 1974. Health Care Management Program, Boston University School of Management. 1976.

Block, L. "Hospitals and Other Institutional Facilities and Services, 1939." Vital Statistics, Special Reports 13, 1-57. Washington, DC: Bureau of the Census, 1942.

Buchanan, J. M. and G. Tullock. The Calculus of Consent, Ann Arbor, Michigan, University of Michigan Press, 1962.

Carroll, D. "The Disability in Hemiplegia Caused by Cerebrovascular Disease." Journal of Chronic Disease. 15:179. February 1962.

Chiswick, Barry R. "The Demand for Nursing Home Care: An Analysis of the Substitution Between Institutional and Non-Institutional Care." Journal of Human Resources. 1976.

Christianson, Jon. "Conflicting Social Goals and the Choice of Regulatory Instruments: An Application in Long-Term Health Care Facilities." Presented at the annual meeting of the Western Economics Association. 1976.

_____. Long Term Care Standards: Enforcement and Compliance. Unpublished paper. Department of Agricultural Economics and Economics, Montana State University. Bozeman, Montana: 1977.

Cohen, Harold. Personal Communication. October 1978.

Community Research Applications. A National Study of Levels of Care in Intermediate Care Facilities. Report for Health Services Administration, prepared under Contract No. HSA 105-74-0176. April 1976.

Comptroller General of the United States. Report to the Congress: Problems in Providing Guidance to States in Establishing Rates of Payment for Nursing Home Care under the Medicaid Program. Department of Health, Education, and Welfare. Washington, DC: April 11, 1975.

Cooper, Barbara S., Nancy L. Worthington and Mary McGee. Compendium of National Health Expenditure Data. Social Security Administration, Office of Research and Statistics, U.S. DHEW. Publication No. 73-11903. 1973.

Courville, Leon. "Regulation and Efficiency in the Electric Utility Industry." Bell Journal. Vol. 5, 1974.

Cromwell, Jerry. "What We Have and Haven't Learned from Prospective Payment Programs." Abt Associates Health Care Systems Area Discussion Paper No. 10. 1977.

Cromwell, Jerry, et al. Analysis of Prospective Payment Systems in Upstate New York. Final Report prepared under Contract No. HEW OS-74-261. Abt Associates Inc. 1976.

Curry, Timothy J., and Bascom W. Ratliff. "The Effects of Nursing Home Size on Resident Isolation and Life Satisfaction." The Gerontologist. Autumn 1973.

Dinnerstein, A.J., M. Lowenthal, and M. Dexter. "Evaluation of a Rating Scale of Ability in Activities of Daily Living." Archives of Physical Medicine and Rehabilitation. Vol. 46: 579. August 1975.

Dowling, William L. "Prospective Rate Setting: Concept and Practice." Topics in Health Care Finance. Vol. 3: 2. Winter 1976.

Downs, Anthomy. An Economic Theory of Democracy. New York: Harper and Row. 1957.

Eagle, Edward. "Nursing Homes and Related Facilities: A Review of the Literature." Public Health Reports, Vol. 83: 8., pp. 673-684. August 1968.

Edwards, Allen L. Techniques of Attitude Scale Construction. Englewood Cliffs: New Jersey: Prentiss Hall, Inc., 1957.

Feldstein, Martin S. "Hospital Cost Inflation: A Study of Nonprofit Price Dynamics." American Economic Review. Vol. 61: no. 5, pp. 853-872. December 1971.

_____. "An Econometric Model of the Medicare System." Quarterly Journal of Economics. Vol. LXXXV: No. 1, pp. 1-20. February 1971.

Fisher, Joseph C. Multi-collinearity and the Quality of Social Science Research. Tufts University, 1977.

Friedman, Bernard. "The Tradeoff Between Delays and Excess Cost in Decisions to Restrict the Supply of Institutional Beds for Long-Term Care." Brown University (mimeographed). April 1975.

Fuller, Wayne A. "Grated Polynomials as Approximating Functions." Australian Journal of Agricultural Economics. June 1969.

General Accounting Office, U.S. Comptroller, General. "Problems in Providing Guidance to States in Establishing Rates of Payment for Nursing Homes Under Medicaid Programs." Report to Congress B-164031(3). April 19, 1972.

Geomet, Inc. "An Evaluation of the Operating Costs and Cost Impacts of Regulations on Nursing Homes in the State of Ohio." Report No. HF-72, prepared for the Federation of Community Planning. Cleveland, Ohio: 1972.

Gersten, J.W., C. Ager, K. Anderson, and F. Cenkovich. "Relation of Muscle Strengths and Range of Motion to Activities of Daily Living." Archives of Physical Medicine and Rehabilitation. Vol. 57: 137. March 1970.

Gibson, Robert M. and Marjorie Smith Mueller. "National Health Expenditures, Fiscal Year 1976." Social Security Bulletin. April 1970.

Ginsburg, Paul B. "Cost Containment Through Reform of Nursing Home Reimbursement." Academic and State Health Policy Issues. Ed. S. Bushkin. Public Services Laboratory. Georgetown University: 1976.

_____. "Problems of Comparing Cost Data." Discussion Paper (mimeograph) prepared for Patient Care Assessment/Cost Seminar. Office of Long Term Care, DHEW. October 6-7, 1976.

Given, Charles W. "A Rationale for Evaluating the Costs and Efficiency of Health Services in Terms of Functional Measures of Patient Outcomes." (Processed 1974.)

Goldfarb, Marsha, Mark Hornbroth, and John Rafferty. "Behavior of the Multiproduct Firm: A Model of the Nonprofit Hospital." Presented at the Western Economics Association Annual Meeting. Anaheim, California: June 21, 1977.

Gordon, E.E. and K. Kuhn. "Evaluation of Rehabilitation Methods in the Hemiplegic Patient." Journal of Chronic Disease. Vol. 19: 3. January 1966.

Gottesman, Leonard E. "Nursing Home Performances as Related to Resident Traits, Ownership, Size, and Source of Payment." American Journal of Public Health. Vol. 64: 3, pp. 269-276. 1974.

Grana, John. "Management Decision Making in the Long-Term Health Care Industry in Massachusetts." University of Massachusetts (processed 1974).

Greenwald, Shayna R. and Margaret W. Linn. "Intercorrelation of Data on Nursing Homes." The Gerontologist. Part 1: pp. 337-340. Winter 1971.

Hamilton, Diane, and Gail Jensen. Health Care Cost Variation: Departmental Analyses Incorporating Case Mix. Abt Associates: October 1, 1978a.

Hamilton, Diane, Michael Sumner, Paul Freddolino, James Tilley, Karen Tyson, and James Boland. Evaluation of the Hospital Reimbursement Experiment for South Carolina Hospitals. Report to HCFA, #SSA 600-76-0047, Abt Associates Inc., 1978b.

Health Care Financing Administration. "Action Transmittal HCFA-AT-77-85, MSB Medical Assistance Manual." Washington, D.C., September 1977.

Hollis, Gloria and Karen Lashman. "Nursing Home Cost Data: A Review of Previous Studies." Prepared for NCHS, U.S. DHEW. (Mimeograph.)

Holmberg, R. Hopkins and Nancy M. Anderson. "Implications of Ownership for Nursing Home Care." Medical Care. Vol. VI: No. 4, pp. 300-307. July-August 1968.

Johnston, J. Econometric Methods. Second edition. New York: McGraw-Hill. 1972.

Jones, Ellen, et al. An Approach to the Assessment of Long Term Care. Final Report No. HS-01162. Harvard Center for Community Health and Medical Care. 1976.

Jorgenson, Dale V., and Sidney S. Handel, "Investment Behavior in the U.S. Regulated Industries." Bell Journal. Vol. 2: 1971.

JWK International Corporation. "Assessment of Cost and Operational Impacts of SNF/ICF Standards." Prepared under Contract No. SRS 500-75-0032 for the Social and Rehabilitation Services, DHEW. 1976.

Kane, Robert L., Lou Ann Jorgensen, and Ginette Pepper. "Can Nursing Home Care be Cost Effective?" Journal of the American Geriatric Society. Vol. XXII: No. 6, pp 265-272.

Katz, S., A.B. Ford, R.W. Moskowitz, B.A. Jackson, and M.W. Jaffe. "The Index of A.D.L.: A Standardized Measure of Biological and Psychosocial Function." Journal of the American Medical Association. Vol. 184: 914. September 21, 1963.

Katz, S., T.D. Downs, H. Cash, and R.C. Grotz. "Progress in Development of the Index of A.D.L." Gerontologist, Vol. 10: 20. Spring 1970.

Keinty, Rita M. "Massachusetts State Health Policy on Long Term Care for the Elderly." Health Policy Study, University of Massachusetts, Institute for Governmental Services. (Processed December 1975.)

Kmenta, Jan. Elements of Econometrics. New York: Macmillan. 1971.

Kosberg, Jordan J. "The Nursing Home: A Social Work Paradox." Social Work. Vol. 18: 2, pp. 104-110. March 1973.

Laudicina, Susan S. Prospective Reimbursement for Hospitals: A Guide for Policy Makers. New York: Community Service Society of New York. 1976.

Lave, Judith R. and Lester B. Lave. Hospital Cost Analysis, Implications for Cost Controls. Final report prepared under Contract No. 106-74-0045 for the National Center for Health Services Research. May 1975.

Lee, M.L. "A Conspicuous Production Theory of Hospital Behavior." Southern Economic Journal. Vol. 38, pp. 48-58. July 1971.

Lee, M.L., and Richard L. Wallace. "Problems in Estimating Multiproduct Cost Functions: An Application to Hospitals." Western Economic Journal. Vol. XI: 3. September 1973.

Levey, Samuel, et al. "An Appraisal of Nursing Home Care." Journal of Gerontology. Vol. 28: 2, pp. 222-228. April 1973.

Linn, M.W. "A Rapid Disability Rating Scale." Journal of the American Geriatric Society. Vol. 15:211. February 1967.

McCoy, Leahmae. "The Nursing Home as a Public Utility." Journal of Economic Issues. Vol. 5: 1, pp. 67-76. March 1971.

Mason, R., R. Gunst, and Webster. "Regression Analysis and Problems of Multicollinearity." Communications in Statistics, Vol. 4, No. 3, 1975.

Medicaid Bureau, Health Care Financing Administration. State Summaries of Current and Proposed Reimbursement Systems. Provided by Milton DeZube, 1978.

Millin, Michail B. "Changing Perspectives and Cross Currents in Nursing Home Care: 1969-1969." Journal of the American Geriatric Society. Vol. 18: 2, pp. 124-141. 1970.

Moreland Commission. Report of the New York State Moreland Commission on Nursing Homes and Residential Facilities. Albany, New York: January 1976.

Moroney, Robert M. and Norman R. Kurtz. "The Evaluation of Long Term Care Institutions." In, Sylvia Sherwood, ed., Long Term Care. New York: Spectrum Publications, pp. 81-121. 1975.

Morris, Robert and Elizabeth Harris. Home Health Services in Massachusetts; 1971: Their Role in the Care of the Long-Term Sick. American Journal of Public Health. Vol. 62: 8, pp. 1088-1093. August 1972.

Mueller, Marjorie Smith and Paula A. Piro. "Private Health Insurance in 1974: A Review of Coverage, Enrollment, and Financial Experience." Social Security Bulletin. Vol. 39, 3, pp. 3-20. March 1976.

Muller, Charlotte, Eleanor Gerber, and Paul Worthington. "Planned Capital Formulation in Nursing Homes in New York City." Inquiry. Vol. VI: 4, pp. 12-26. December 1969.

Muller, J.N., J.S. Tobias and H.R. Kelman. "Rehabilitation Potential of Nursing Home Residents." American Journal of Public Health. Vol. 53: 243. February 1963.

Muller, Lawrence, R.P. Johnson and K.M. McCaffree. Economic Analysis of the Production Process in Washington State Nursing Homes. Preliminary Draft Report. Batelle Memorial Institute. (Processed 1975.)

Mundlak, Yair. "Transcendental Multiproduct Production Functions." International and Economic Review. Vol. 5: 3. September 1964.

Murnaghan. Jane H. (editor). "Long Term Care Data." Report of the Conference on Long Term Health Care Data. Medical Care. Vol. 14:5, supplement. May 1976.

National Council of Health Care Services. "A Payment System for Long-Term Care Institutional Services." 1977.

National Center for Health Satistics, U.S. DHEW. Selected Operating and Financial Characteristics of Nursing Homes in the United States: 1973-1974 National Nursing Home Survey. DHEW Publication No. HRA 76-1773. Rockville, Maryland: Health Resources Administration. December 1975.

_____. Nursing Homes in the United States: 1973-1974. Vital and Health Statistics. Series 14: 17. October 1977.

Nerlove, Marc. Estimation and Identification of Cobb-Douglas Production Functions. Amsterdam, North Holland: 1965.

Newhouse, Joseph P. "Toward a Theory of Nonprofit Institutions: An Economic Model of a Hospital." American Economic Review. Vol. 60: No. 1, pp. 64-74. May 1970.

Niskanen, William A. Bureaucracy and Representative Government. Aldine Press, Chicago, 1971.

Noll, Roger. "The Consequences of Public Utility Regulation of Hospitals." Sun Valley Conference on Health Care Regulation.

Nunnally, Jum C. Psychometric Theory. New York: McGraw Hill, Inc. 1967.

Orzechowski, William. "Economic Models of Bureaucracy: Survey, Extensions, and Evidence." In Budgets and Bureaucracies: The Sources of Government Growth. Ed. by Thomas E. Borcherding. Duke University Press, 1977.

Pauly, Mark and M. Redisch. "The Not-for-Profit Hospital as a Physician's Cooperative." American Economic Review. Vol. 6: 1, pp. 87-99. March 1973.

Petersen, H. Craig. "An Empirical Test of Regulatory Effects." Bell Journal. Vol. 6. Spring 1975.

Pfouts, Ralph W. "The Theory of Cost and Production in the Multi-Product Firm." Econometrics. Vol. 29: 4. October 1961.

Pindyck, R. S. and D. C. Rubinfeld. Econometric Models and Economic Forecasts. New York: McGraw-Hill, 1976.

Reinhardt, Uwe. Physician Productivity and the Demand for Health Manpower. Cambridge: Ballinger Publishing Co. 1975.

Robinson, Ira E., Harold Bronfin, and Jack O. Balswick. "Nursing Home Resources and Rural and Urban Needs." Journal of Health and Social Behavior. Vol. 11, pp. 146-151. June 1971.

Ruchlin, Hirsch S. "A New Strategy for Regulating Long-Term Care Facilities." Health, Politics, Policy, and Law. Summer, 1977.

Ruchlin, Hirsch S. and Samuel Levey. "Nursing Home Cost Analysis: A Case Study." Inquiry. Vol. IX: 3. 1972.

_____. "An Economic Perspective of Long Term Care." In Ed. Sylvia Sherwood, Long Term Care. New York: Spectrum Publications, pp. 125-177. 1975.

Ruchlin, Hirsch S., Samuel Levey and Charlotte Muller. "The Long Term Care Marketplace: An Analysis of Deficiencies and Potential Reform by Means of Incentive Reimbursement.

_____. "An Incentive Reimbursement Proposal for Financing Long Term Care." Final Report prepared under Contract No. HEW OS 74-176. September 30, 1974.

Samuelson, Paul A. Foundations of Economic Analysis. Cambridge, Mass: Harvard University Press. 1975.

Sattler, Fredric L., ed. Hospital Prospective Payment. Interstudy. Mineapolis, Minnesota: 1976.

Schelling, Thomas C. The Strategy of Conflict. Cambridge: Harvard University Press. 1960.

Schoening, H.A. and K.A. Iversen. "Numerical Scoring of Self-Care Status: A Study of the Kenny Self-Care Evaluation." Archives of Physical Medicine and Rehabilitation. Vol. 49: 221. April 1968.

Schulman, David and Ruth Galanter. "Reorganizing the Nursing Home Industry: A Proposal." Milbank Memorial Fund Quarterly. Vol. 54: 2, pp. 129-143. Spring 1976.

Schultze, Charles." The Public Use of Private Interest." Harper's Magazine. May 1977.

Scranton, J., M.L. Fogel, and W.J. Erdman. "Evaluation of Functional Levels of Patients During and Following Rehabilitation. Vol. 51:3. January 1970.

Shepard, Roger, Kimball Romney, and Sara Beth Nerlove. Multidimensional Scaling: Theory and Application in the Behavioral Sciences. New York: Siminar Press. 1972.

Singler, George. "A Theory of Regulation." Bell Journal. Spring 1971.

Skinner, Douglas E. and Donald E. Yett. "Debility Index for Long-Term Care Patients." In Robert L. Berk (ed.). Health Status Indices. Chicago: Hospital Research and Educational Trust. 1973.

_____. "Estimation of Cost Functions for Health Services: The Nursing Home Case." Paper presented to SEA. Human Resources Research Center, University of Southern California Grant 10-P-56030/9-04, SRS, 1970.

Sloan, Frank A. and Bruce Steinwald. "Effects of Regulation on Hospital Costs and Input Use." Presented at the American Economic Association Conference. Chicago, Illinois: August 29, 1978.

Sloss, James. "Regulation of Motor Freight Transportation: A Quantitative Evaluation of Policy." Bell Journal. Vol. 1. 1970.

Social and Rehabilitation Services, DHEW. "Information Relating to Payment Rates and Methods Followed in Establishing Payment Rates for Title XIX Skilled Nursing Homes and Intermediate Care Facilities." Unpublished survey questionnaire data made available by Betty Cornelius. 1973.

Solon, Jerry Alan. "Medical Care: Its Social and Organizational Aspects, Nursing Homes and Medical Care." New England Journal of Medicine. Vol. 269: 20. November 14, 1963.

Solon, Jerry Alan and A.M. Baney. "Inventory of Nursing Homes and Related Facilities." Public Health Reports. Vol. 69, pp. 1121-1132. December 1954.

Spann, Robert M. "Rate of Return Regulation and Efficiency in Production: An Empirical Test of the Anerch-Johnson Thesis." Bell Journal. Vol. 5. 1974.

Theil, Henri. Principles of Econometrics. New York: John Wiley and Sons, Inc. 1971.

Touche Ross and Company. "A Study of Cost Differentials Between Skilled Nursing Facilities and Intermediate Care Facilities." Report to the U.S. Department of Health, Education and Welfare. (Processed January 31, 1974.)

Urban Institute. Literature on nursing home regulations provided to Howard Birnbaum of Abt Associates Inc. by William Scalon.

U.S. Census. Statistical Abstract of the United States: 1977. United States Government Printing Office. Washington, DC: 1977.

U.S. Congress, Special Subcommittee on Aging, United States Senate. Nursing Home Care in the United States: Failure in Public Policy. Washington, DC: Government Printing Office. 1974.

U.S. DHEW, Social and Rehabilitation Services. "Information Relating to Payment Rates and Methods Followed in Establishing Payment Rates for Title XIX Skilled Nursing Homes and Intermediate Care Facilities." Unpublished survey questionnaire data (1973) made available by Betty Cornelius.

U.S. DHEW. Trends; 1965 Edition. Part 1, National Trends, S-96.

_____. Compendium of National Health Expenditures Data. DHEW Publication No. (SSA) 73-11903, Washington, DC: 1972. Government Printing Office.

_____. Medicare Provider Reimbursement Manual. Social Security Administration. Government Printing Office Reprint No. HIM 15(7-67). Washington, DC: 1976.

_____. Selected Operating and Financial Characteristics of Nursing Homes. Public Health Service, Vital and Health Statistics, Series 13, No. 2, publication No. (HRA) 76-1773. December 1975.

U.S. Department of Labor. Handbook of Labor Statistics 1975--Reference Edition. U.S. Government Printing Office, Bulletin 1865. 1975.

Wallace, T.D. and Ashig Hussain. Econometrica. Vol. 37: 1. January 1969.

Wallis, Kenneth. Topics in Applied Econometrics. London: Grey-Mills Publishing Ltd. 1972.

Walters, A. "Production and Cost Functions: An Econometric Survey," Econometrica, Vol. 31, No. 1-2 (January-April, 1963).

Wiles, P. J. D. Price, Costs, and Output. New York: Frederick A. Praeger. 1961.

Winn, Sharon. "Assessment of Cost Related Characteristics and Conditions of Long-Term Care Patients." Inquiry. December 1975.

_____. "Analysis of Selected Characteristics of a Matched Sample of Nursing Homes." Medical Care. March 1974.

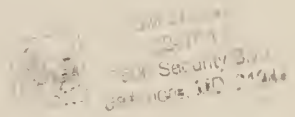
Winn, Sharon, Carol Bennett, and Kenneth M. McCaffree. "The Relationship of Resident Patient Condition and Characteristics to Variations in Costs of Long Term Care." Paper presented at American College of Nursing Home Administrators. September 17, 1976.

Walsh, Thomas J. "Service Unit Measurement and the Pace Data Collection Form." Illinois Department of Public Health; discussion Paper for Patient Care Assessment Cost Seminar, DHEW. October 6-7, 1976.

Wylie, C.M. "Administrative Research in the Rehabilitation of Stroke Patients." Rehabilitation Literature. Vol. 25. January 1964.

Young, J.P. "A Method for Allocation of Nursing Personnel to Meet Inpatient Care Needs." Operations Research Division, The Johns Hopkins Hospital. October 1962.

Zimmer, James J. "Characteristics of Patients and Care Provided in Health Related and Skilled Nursing Facilities." Medical Care. Vol. 13: 12, pp. 992-1010. December 1975.



CMS LIBRARY
3 8095 00007950 5